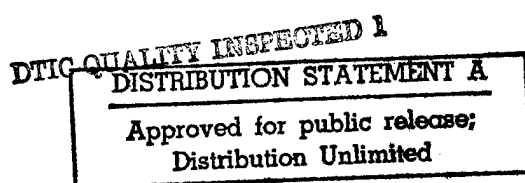


**An Evaluation of the Effectiveness of the  
Polygraph Automated Scoring System (PASS)  
in Detecting Deception in a  
Mock Crime Analog Study**

**N. Joan Blackwell, M.S.  
March 1994**

Department of Defense Polygraph Institute  
Fort McClellan, Alabama 36205-5114  
Telephone: 205-848-3803  
FAX: 205-848-5332

19960326 003



Encl 2

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE March 1994	3. REPORT TYPE AND DATES COVERED Final Report (From May 93 - March 94)		
4. TITLE AND SUBTITLE An Evaluation of the Effectiveness of the Polygraph Automated Scoring System (PASS) in Detecting Deception in a Mock Crime Analog Study		5. FUNDING NUMBERS DoDPI94-P-0023		
6. AUTHOR(S) N. Joan Blackwell				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Department of Defense Polygraph Institute Building 3165 Fort McClellan, AL 36205-5114		8. PERFORMING ORGANIZATION REPORT NUMBER DoDPI94-R-0003		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Department of Defense Polygraph Institute Building 3165 Fort McClellan, AL 36205-5114		10. SPONSORING/MONITORING AGENCY REPORT NUMBER DoDPI94-P-0023 DoDPI94-R-0003		
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT  Public release, distribution unlimited.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The purpose of this research was to evaluate the effectiveness of a prototype scoring algorithm, the Polygraph Automated Scoring System (PASS), version 2.0, when analyzing mock crime data. A database of psychophysiological detection of deception (PDD) measurements was collected using the Axciton Polygraph System. The PDD test format used was the Zone Comparison Test (ZCT). In addition, two types of control question tests (CQTs) were used during the study: (a) an experimental version of the directed lie control (DLC), and (b) the conventional probable lie control (PLC) currently in use throughout the PDD community. The data set is composed of 120 local civilian citizens who underwent a PDD examination, half of which committed a mock theft of money. The results showed that PASS accuracy on controlled laboratory data with known ground truth was far below that attained by APL during algorithm development when using "live" PDD examinations. In fact, rates of accuracy for both PASS and the PDD examiners were lower than anticipated. PASS was clearly more accurate in identifying individuals who had been programmed "innocent," while the PDD examiners were more accurate in detecting individuals who had been programmed "guilty." Use of the DLC versus the PLC made no apparent difference in overall accuracy for either PASS or the PDD examiners, though there were statistically significant differences when analyzed from the standpoint of programmed condition. Analysis also showed a PDD examiner performance decrement over time which contributed to the lower than expected examiner accuracy rates.				
14. SUBJECT TERMS Polygraph Automated Scoring System (PASS), Axciton, computerized scoring algorithms, probable lie control test format, directed lie control test format, control question test, Zone Comparison Test, polygraph, forensic (PDD)			15. NUMBER OF PAGES 81	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT	

Report No. DODPI94-R-0003

An Evaluation of the Effectiveness of the  
Polygraph Automated Scoring System (PASS)  
in Detecting Deception in a  
Mock Crime Analog Study

N. Joan Blackwell, M.S.

March 1994

Department of Defense Polygraph Institute  
Fort McClellan, Alabama 36205

## Director's Foreword

This report describes a study designed to evaluate the relative accuracy of examiner and computer algorithm scoring of psychophysiological detection of deception (PDD) examinations. It is important, if not essential, that this type of cross validation study be completed to assess the validity of computer algorithms designed to evaluate PDD examinations. Results of this study demonstrate that there are no significant differences between overall examiner and computer algorithm accuracies, suggesting that computer algorithm scoring can be as accurate as human scoring. It should be noted that the reported comparison was made using data collected following a laboratory mock-crime while the algorithm was designed using the results of actual criminal examinations. This may have reduced overall accuracy rates if there are intrinsic differences between data collected following actual and mock crimes.

A handwritten signature in dark ink, reading "Michael H. Capps". The signature is fluid and cursive, with the first name "Michael" being the most prominent part.

Michael H. Capps  
Director

## Acknowledgments

The author wishes to express appreciation to the following organizations for providing manpower and resources in support of this research: the Department of Defense Polygraph Institute, (DoDPI); the Air Force Office of Special Investigations, (AFOSI); the Naval Investigative Service (NIS); the Defense Investigative Service, (DIS). Special thanks are extended to the individuals who performed tasks associated with project coordination and execution.

SA Dan V. Weatherman (AFOSI) and SA Robert P. Fritzsche (NIS) conducted all psychophysiological detection of deception (PDD) examinations. SA Robert P. Cole (DIS) and SA David L. Miller (DIS) performed all blind scoring of the PDD examinations.

Charlene L. Stephens (DoDPI) served as research assistant during data collection, assisted in setting up the study files and generated a seminar briefing packet on the preliminary data. Sarah E. Tidwell (DoDPI) served as the "deliberate intruder" at the crime scene and assisted in setting up the study files. Jeffrey E. St. Cyr (DoDPI) entered the demographic and evaluation data into a statistical analysis computer file and created props used during the scenario. Joan Harrison-Woodard (DoDPI) assisted in generating a seminar briefing packet on the preliminary data. Brenda J. Smith (DoDPI) was instrumental in manpower and resource coordination throughout the project.

Karen M. Molloy (DoDPI) interfaced with the employment agency on a daily basis to resolve examinee scheduling problems. Frank M. Ragan (DoDPI) ensured the availability of computer disks, video tapes, and other required supplies. MSGT Randall S. Reynolds (AFOSI) set up and maintained the Axciton equipment and video recording equipment. Frederick H. Fisher (DoDPI) seated the examinees and announced their arrival to the research staff. Genni B. Arledge (DoDPI) ensured that the examinees remained in the post examination holding area until being released by the researcher.

This study was supported by funds from the DODPI as project DoDPI93-P-0023. The views expressed in this article are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

## Abstract

BLACKWELL, N. J. An Evaluation of the Effectiveness of the Polygraph Automated Scoring System (PASS) in Detecting Deception in a Mock Crime Analog Study. March 1994, Report No. DoDPI94-R-0003 Department of Defense Polygraph Institute, Fort McClellan, AL 36205.--The purpose of this research was to evaluate the effectiveness of a prototype scoring algorithm, the Polygraph Automated Scoring System (PASS), Version 2.0, when analyzing mock crime data. A database of psychophysiological detection of deception (PDD) measurements was collected using the Axciton Polygraph System. The PDD test format used was the Zone Comparison Test (ZCT). In addition, two types of control question tests (CQTs) were used during the study: (a) an experimental version of the directed lie control (DLC), and (b) the conventional probable lie control (PLC) currently in use throughout the PDD community. The data set is composed of 120 local civilian citizens who underwent a PDD examination, half of which committed a mock theft of money. The results showed that PASS accuracy on controlled laboratory data with known ground truth was far below that attained by APL during algorithm development when using "live" PDD examinations. In fact, rates of accuracy for both PASS and the PDD examiners were lower than anticipated. PASS was clearly more accurate in identifying individuals who had been programmed "innocent," while the PDD examiners were more accurate in detecting individuals who had been programmed "guilty." Use of the DLC versus the PLC made no apparent difference in overall accuracy for either PASS or the PDD examiners, though there were statistically significant differences when analyzed from the standpoint of programmed condition. Analysis also showed a PDD examiner performance decrement over time which contributed to the lower than expected examiner accuracy rates.

Key-words: Polygraph Automated Scoring System (PASS), Axciton, computerized scoring algorithms, probable lie control (PLC) test format, directed lie control (DLC) test format, control question test (CQT), Zone Comparison Test (ZCT), polygraph, forensic psychophysiology, psychophysiological detection of deception (PDD)

## Executive Summary

BLACKWELL, N. J. An Evaluation of the Effectiveness of the Polygraph Automated Scoring System (PASS) in Detecting Deception in a Mock Crime Analog Study. March 1994, Report No. DoDPI94-R-0003 Department of Defense Polygraph Institute, Fort McClellan, AL 36205.

The prototype scoring algorithm, Polygraph Automated Scoring System (PASS), Version 2.0, developed by the Johns Hopkins University Applied Physics Laboratory (APL) was reported to have an overall accuracy rate of 99.4% on the 374 field cases used during development. Additionally, for a subset of 91 cases which had confirmed decisions the accuracy rate was cited as 100.0%.

The purpose of this research was to evaluate the effectiveness of PASS when analyzing mock crime data. Comparisons were also made between PASS accuracy and the psychophysiological detection of deception (PDD) examiners who conducted the examinations.

### Procedure

A database of 120 PDD examinations was collected using the Zone Comparison Test (ZCT) in conjunction with the Axciton Polygraph System. Two types of control question tests (CQTs) were used during the study: (a) an experimental version of the directed lie control (DLC), and (b) the probable lie control (PLC) currently in use throughout the PDD community.

Examinees for this study were local civilian citizens provided by a contracted employment service. Half of the examinees committed a mock theft of money (\$124.00) from the "Country Store" and were instructed to lie about their involvement in the crime. A scenario participant known as the "deliberate intruder" interrupted the theft in progress, forcing the examinee to conceal his/her actions; thereby heightening the individual's level of arousal.

### Findings

The results showed that the PASS was more accurate in identifying individuals who had been programmed "innocent," while the PDD examiners were more accurate in detecting individuals who had been programmed "guilty." However, overall accuracy rates were lower than expected for both PASS and the PDD examiners. Further analysis showed a PDD examiner performance decrement over time which contributed to the lower than expected accuracy rates. Use of the DLC versus the PLC made no apparent difference in overall accuracy for either PASS or the PDD examiners, though there were statistically significant differences when analyzed from the standpoint of programmed condition.

A subset of data (73 cases), which met one of the "ground truth" criteria used by the APL during PASS development was also analyzed. Of 26 cases labeled no deception indicated (NDI) by both the PDD examiners and the blind scorers, PASS agreed with their decision on 100.0% of the cases. (This resulted in 23 correct calls and three false negative [FN] decisions when compared to ground truth.) Of the 45 cases judged as deception indicated (DI) by the group of examiners, PASS agreed with their decision on 64.4% of the cases. (In this comparison, the PDD examiners generated 37 correct DI decisions and 8 false positives [FPs]. PASS generated 29 correct DI decisions, 5 FPs, and 11 INCs.) PASS accuracy on these DI cases was far below that attained by APL during algorithm development.

Overall PASS provided moderate levels of accuracy when coupled with the decisions of the original examiner and two blind scorers. It was less accurate when used alone. Additionally, the inconclusive (INC) rate generated by PASS was usually much higher than the rate produced by the PDD examiners.

#### Utilization of Findings

The findings in this report provide PDD managers with a means to assess the potential usefulness of a scoring system, such as PASS, for their respective organizations. The prototype, PASS Version 2.0, is a user friendly software package, however the data collected during this study did not support the level of accuracy reported by APL during development. Use of a computerized system such as PASS offers the opportunity for scoring consistency. However, since PASS enables the examiner to edit perceived artifacts, scoring consistency will be ensured only if all reviewing examiners edit the same tests in the same manner. Another consideration for management is the costs associated with retesting individuals judged INC by PASS.



## Table of Contents

Title page . . . . .	i
Director's Foreword . . . . .	ii
Acknowledgements . . . . .	iii
Abstract . . . . .	iv
Executive Summary . . . . .	v
List of Figures . . . . .	ix
List of Tables . . . . .	x
Introduction . . . . .	1
Method . . . . .	7
Research Design . . . . .	7
Subjects . . . . .	9
PDD Examiners . . . . .	9
Apparatus . . . . .	11
Hardware . . . . .	11
Software . . . . .	11
Crime Scene . . . . .	12
Scenario . . . . .	12
Innocent Subjects . . . . .	12
Guilty Subjects . . . . .	12
Procedure . . . . .	13
Scoring Criteria and Data Analysis Procedures . . . . .	15
PDD Examiner . . . . .	15
PASS . . . . .	15
Results . . . . .	16
Overall Accuracy Comparisons . . . . .	17
PASS Decision vs. Ground Truth . . . . .	17
PDD Examiner Decision vs. Ground Truth . . . . .	18
Chi-Square and Fisher Exact Statistics . . . . .	18
NDI and DI Accuracy Comparisons . . . . .	19
PASS Decision vs. Ground Truth . . . . .	19
PDD Examiner Decision vs. Ground Truth . . . . .	19
Chi-Square and Fisher Exact Statistics . . . . .	20
Control Question Test (CQT) Comparisons . . . . .	21
PASS Decision vs. Directed Lie Control (DLC) . . . . .	21
PASS Decision vs. Probable Lie Control (PLC) . . . . .	22
Chi-Square and Fisher Exact Statistics . . . . .	23
PDD Examiner Decision vs. DLC . . . . .	23
PDD Examiner Decision vs. PLC . . . . .	25
Chi-Square and Fisher Exact Statistics . . . . .	26
PASS/PDD Examiner Comparisons Using DLC/PLC Cases . . . . .	27
Chi-Square and Fisher Exact Statistics . . . . .	28
Interrater Agreement Analyses . . . . .	29
Blind Scoring . . . . .	29
PDD Examiner/Blind Scorer Decision vs. PASS Decision . . . . .	30
PASS/PDD Examiner/Blind Scorer Decision vs. Ground Truth . . . . .	31
Accuracy Compared Over Time . . . . .	32
PDD Examiner Accuracy . . . . .	32
PASS Accuracy . . . . .	33
PASS and PDD Examiner Accuracy . . . . .	34

## Table of Contents, cont.

Discussion . . . . .	35
Accuracy . . . . .	35
Interrater Agreement . . . . .	37
Defense of the Findings . . . . .	38
Effect of Medications . . . . .	38
Prior Criminal Record of Examinee . . . . .	38
Strength of the Mock Crime Scenario . . . . .	39
Examiner Performance Decrements Over Time . . . . .	39
Artifact Editing . . . . .	40
PASS Accuracy via APL Assessment Criteria . . . . .	40
Conclusions . . . . .	41
References . . . . .	43
Appendix A: Chart Evaluation Criteria . . . . .	A-1
Appendix B: "Programmed Innocent" Script . . . . .	B-1
Appendix C: "Programmed Guilty" Script . . . . .	C-1
Appendix D: "Crime" Scene Scenario Questionnaire . . . . .	D-1
Appendix E: Welcome and General Briefing . . . . .	E-1
Appendix F: Project Briefing . . . . .	F-1
Appendix G: Volunteer Agreement Affidavit . . . . .	G-1
Appendix H: Background Information Form . . . . .	H-1
Appendix I: Directed Lie Control (DLC) Pretest Interview . . . . .	I-1
Appendix J: Probable Lie Control (PLC) Pretest Interview . . . . .	J-1
Appendix K: Rights Advisement Form . . . . .	K-1
Appendix L: Consent to Interview with Polygraph . . . . .	L-1
Appendix M: Stimulation (numbers) Test . . . . .	M-1
Appendix N: Zone Comparison Test Question List . . . . .	N-1
Appendix O: Subject Debriefing Form . . . . .	O-1

## List of Figures

1. Response intervals, or scoring windows for the four PDD channels generated by Polygraph Automated Scoring System (PASS), Version 2.0 . . . . .	2
2. Diagram showing experimental design. . . . .	8
3. Diagram showing examiner collection sequence . . . . .	8
4. Flow chart of principal steps in data collection process	15
5. Diagram showing Zone Comparison Test (ZCT) question sequence and evaluation spots with compared control and relevant questions linked . . . . .	16
6. Chart showing numerical evaluation criteria for Zone Comparison Test (ZCT) . . . . .	16

## List of Tables

1. PASS Signal Scoring Weights . . . . .	2
2. Percentage of Agreement Between PASS and PDD Examiner . .	4
3. Percentage of Agreement Between PASS Decisions and Ground Truth . . . . .	4
4. Percentage of Agreement Between PASS and PDD Examiner on NDI and DI Decisions During Algorithm Development . . . .	5
5. Subject Demographics . . . . .	10
6. Physical Condition of Subjects . . . . .	10
7. Data Excluded from Analysis . . . . .	10
8. PDD Examiner Qualifications . . . . .	11
9. Percentage of Agreement Between PASS Decision and Ground Truth . . . . .	17
10. Percentage of Agreement Between PDD Examiner Decision and Ground Truth--All Cases . . . . .	18
11. Result of Chi-square ( $X^2$ ) and Fisher Exact Statistics when Testing Association Between PASS and PDD Examiner Decisions--All Cases . . . . .	19
12. Percentage of Agreement Between PASS Decision and Ground Truth--Programmed Innocent and Programmed Guilty Cases .	20
13. Percentage of Agreement Between PDD Examiner Decision and Ground Truth--Programmed Innocent and Programmed Guilty Cases . . . . .	20
14. Results of Chi-square ( $X^2$ ) and Fisher Exact Statistics when Testing Association Between PASS and PDD Examiner Decisions--Programmed Innocent and Programmed Guilty Cases . . . . .	21
15. Percentage of Agreement Between PASS Decision and Ground Truth--All Directed Lie Control (DLC) Cases . . .	22
16. Percentage of Agreement Between PASS Decision and Ground Truth--Programmed Innocent and Programmed Guilty Cases using Directed Lie Control (DLC) . . . . .	22
17. Percentage of Agreement Between PASS Decision and Ground Truth--All Probable Lie Control (PLC) Cases . . .	23

# List of Tables, cont.

18. Percentage of Agreement Between PASS Decision and Ground Truth--Programmed Innocent and Programmed Guilty Cases using Probable Lie Control (PLC) . . . . .	24
19. Results of Chi-square ( $X^2$ ) and Fisher Exact Statistics when Testing Association Between Directed Lie Control (DLC) and Probable Lie Control (PLC) Cases--PASS Decisions . . . . .	24
20. Percentage of Agreement Between PDD Examiner Decision and Ground Truth--All Directed Lie Control (DLC) Cases .	25
21. Percentage of Agreement Between PDD Examiner Decision and Ground Truth--Programmed Innocent and Programmed Guilty Cases using Directed Lie Control (DLC) . . . . .	25
22. Percentage of Agreement Between PDD Examiner Decision and Ground Truth--All Probable Lie Control (PLC) Cases .	26
23. Percentage of Agreement Between PDD Examiner Decision and Ground Truth--Programmed Innocent and Programmed Guilty Cases using Probable Lie Control (PLC) . . . . .	26
24. Results of Chi-square ( $X^2$ ) and Fisher Exact Statistics when Testing Association Between Directed Lie Control (DLC) and Probable Lie Control (PLC) Cases--PDD Examiner Decisions . . . . .	27
25. Results of Chi-square ( $X^2$ ) and Fisher Exact Statistics when Testing Association Between PASS and PDD Examiner--Programmed Innocent and Programmed Guilty Cases using Directed Lie Control (DLC). . . . .	28
26. Results of Chi-square ( $X^2$ ) and Fisher Exact Statistics when Testing Association Between PASS and PDD Examiner--Programmed Innocent and Programmed Guilty Cases using Probable Lie Control (PLC). . . . .	29
27. Kappa Statistics for Blind Scoring Comparisons . . . . .	30
28. Percentage of Agreement Between PASS Decision and the Consensus Decision of the PDD Examiner and Two Blind Scorers . . . . .	31
29. Percentage of Agreement Between Ground Truth and the Consensus Decision of PASS, the PDD Examiner and Two Blind Scorers . . . . .	32
30. PDD Examiner Accuracy Compared Over Time . . . . .	33

List of Tables, cont.

31. PASS Accuracy Compared Over Time . . . . .	34
32. PASS and Examiner One Accuracy Compared Over Time-- Programmed Innocent and Programmed Guilty Cases . . . . .	34
33. PASS and Examiner Two Accuracy Compared Over Time-- Programmed Innocent and Programmed Guilty Cases . . . . .	35

A diagnostic technique which relies on human interpretation of test data is immediately suspect (Nunnally, 1978). Rater bias, inexperience and even incompetence are problems that plague any field in which humans are asked to make interpretive judgements. For more than fifty years the data resulting from psychophysiological detection of deception (PDD) examinations has essentially relied on human interpretation. Accordingly, much of the scientific community considers such data suspect. That contention, along with the ever present need to accurately decipher the complex physiological tracings generated during a PDD exam, are the driving forces behind the development of automated, algorithm-based scoring systems. This research effort was undertaken to evaluate the prototype of one such system.

The most recent endeavor to eliminate subjectivity from the process of interpreting PDD exams is the Polygraph Automated Scoring System (PASS). PASS, a personal computer software package, implements a scoring algorithm developed by the Johns Hopkins University Applied Physics Laboratory (APL) under contract to the National Security Agency (NSA).

The algorithm uses a logistic regression model, and during processing the data is detrended, mathematically filtered and then standardized. PASS currently works in conjunction with the Axciton Computerized Polygraph, a stand alone PDD system developed and marketed by Axciton Systems, Incorporated, Houston, TX. Axciton records the physiological data (i.e., respiration, electrodermal and cardio) collected during a PDD examination. PASS then, in turn, uses that physiological data to produce an overall probability of deception for the examination (Johns Hopkins University Applied Physics Lab, 1993a.)

The scoring criteria currently taught at the Department of Defense Polygraph Institute (DoDPI) was used as a starting point in the development of the PASS algorithm, (Capps, 1993). Along with other selected criteria, APL used the list of factors (see Appendix A) taught in the Chart Evaluation portion of the Basic Courses in Forensic Psychophysiology curriculum (DoDPI, 1992) as a basis to generate approximately 1500 "features" for analysis.

This was accomplished by creating combinations of factors along with varying the number of seconds in the scoring window. Systematically, the list was distilled to include only those features, or criteria which contributed to the highest level of accuracy when used in PDD examination evaluation (Capps, 1993). As a result, expanded scoring windows were established for each of the PDD components, and the cardio channel was split into a pulse channel and a blood volume rate of change (derivative) channel (see Figure 1). The physiological signals were also assigned scoring weights as shown in Table 1.

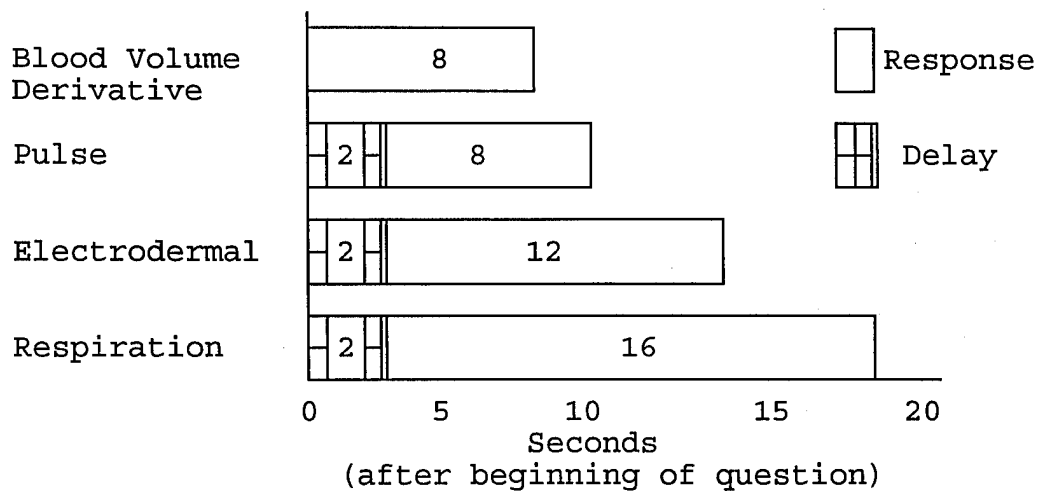


Figure 1. Response intervals, or scoring windows for the four PDD channels generated by Polygraph Automated Scoring System (PASS) Version 2.0. From "Polygraph Automated Scoring System, Version 2.0" by M. Capps, 1993. Adapted by permission.

Table 1  
PASS Signal Scoring Weights

Channel	Percent
Blood Volume Derivative	21
Pulse	14
Electrodermal	49
Respiration	16
Total	100

Note. From "Polygraph Automated Scoring System, Version 2.0" by M. Capps, 1993. Adapted by permission. PASS = Polygraph Automated Scoring System.

Unable to produce acceptable results using mock crime data during the early stages of the algorithm's development, APL ultimately decided to use "live" examinations collected by various PDD agencies. In fact, field cases were used, both to develop the algorithm, and also to later assess its accuracy. Use of field examinations rather than the laboratory-generated mock crime data presented a distinct problem, however; the ground truth information necessary for accuracy assessments was not readily available in the field examinations.

As a result, APL established a two-component guideline to aid in algorithm development. The first component allowed for the incorporation of cases which had been resolved, either through the confession of the examinee or someone else. The second



component enabled the algorithm developers to include cases which, when evaluated, had been assigned the same decision by three qualified examiners--the original examiner and two other examiners who had been designated to blind score the tests (Capps, 1993). Therefore examinations judged either deception indicated (DI), no deception indicated (NDI), or inconclusive (INC) were used during algorithm development, provided all three examiners had arrived at the same decision.

PASS's level of accuracy was initially defined as its rate of agreement with the combined decisions from both the resolved cases and the cases evaluated by the three examiners. It is important to note, however, that as the details of various case resolutions filtered in from the field examiners, the rate of agreement with confirmed ground truth continued to be factored into the analyses.

Of the 374 cases, or subjects used to develop the PASS prototype (Version 2.0), the probability of deception generated by the algorithm supported ground truth (i.e., actual case resolution or the decision of the three examiners) on 93.3% of them, disagreed on 0.5%, and resulted in 6.2% of the cases erroneously being labeled INC (see Table 2). That is, using one of the two methods for determining ground truth, the developers labeled each case either DI, NDI or INC and the algorithm agreed with the respective decision on 349 cases and disagreed twice. On the remaining 23 cases (all deemed to be either DI or NDI), a decision probability of INC was generated.

When APL eliminated those 23 cases from the analysis (as would be done in PDD field accuracy reporting) PASS's rate of agreement with ground truth was 99.4%, and the rate of disagreement was 0.6% (Capps, 1993). The above calculations were based on an internal PASS cut score of 80-20 (i.e., any examination receiving a probability score of 0.80 or higher was considered DI, and any score of 0.20 or lower was considered NDI, while scores falling somewhere in between were termed INC).

As mentioned earlier, ground truth is being provided to APL by the field examiners when each case has been resolved. In early 1992, when PASS Version 2.0 was distributed for use in research efforts such as the one described in this report, 91 of the 374 cases had been provided. In that subset of confirmed data, 93.5% of the cases were correctly identified by PASS, with 6.6% having been labeled INC (see Table 3). When APL removed the inconclusive decisions, PASS provided 100.0% accuracy (Capps, 1993).

Table 2  
Percentage of Agreement Between PASS and "Ground Truth" (N = 374)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Agreed	93.3	(349)	99.4	(349)
Disagreed	0.5	(2)	0.6	(2)
Inconclusive	6.2	(23)	--	--

Note. From "Polygraph Automated Scoring System, Version 2.0" by M. Capps, 1993. Adapted by permission. PASS = Polygraph Automated Scoring System.

\* actual case resolution or the decision of the original examiner and two blind scorers.

Table 3  
Percentage of Agreement Between PASS Decisions and Ground Truth--All Confirmed Cases (N = 91)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Agreed	93.5	(85)	100.0	(85)
Disagreed (FP + FN)	0.0	(0)	0.0	(0)
Inconclusive	6.6	(6)	--	--

Note. From "Polygraph Automated Scoring System, Version 2.0" by M. Capps, 1993. Adapted by permission. FN = false negative; FP = false positive; PASS = Polygraph Automated Scoring System.

A further breakdown of the data, from the standpoint of NDI calls versus DI calls indicated that PASS agreed with the examiners on 94.1% of the cases evaluated as NDI, and judged 5.9% as inconclusive (see Table 4). When APL eliminated the INC calls, PASS provided 100.0% accuracy on NDI decisions.

Analysis of the cases with DI decisions revealed that 92.9% had been identified by PASS as DI, with a 1.1% false negative rate, and another 6.1% of the cases having resulted in a call of inconclusive. When APL removed the INC calls, PASS provided 98.8% accuracy with a 1.2% false negative rate on the DI cases.

Table 4  
Percentage of Agreement Between PASS and "Ground Truth"<sup>a</sup> on NDI and DI Decisions During Algorithm Development (N = 374)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
NDI ( <u>n</u> = 202)				
Agreed	94.1	(190)	100.0	(190)
Disagreed (FP)	0.0	(0)	0.0	(0)
Inconclusive	5.9	(12)	--	--
DI ( <u>n</u> = 182)				
Agreed	92.9	(169)	98.8	(169)
Disagreed (FN)	1.1	(2)	1.2	(2)
Inconclusive	6.1	(11)	--	--

Note. From "Polygraph Automated Scoring System, Version 2.0" by M. Capps, 1993. Adapted by permission. DI = deception indicated; FN = false negative; FP = false positive; NDI = no deception indicated; PASS = Polygraph Automated Scoring System.

<sup>a</sup> actual case resolution or the decision of the original examiner and two blind scorers.

The PASS Version 2.0 has been designed to score a specific type of PDD examination known as the Zone Comparison Test (ZCT). Developed by Mr. Cleve Backster, the ZCT format was adopted and modified by the U. S. Army Military Police School in 1961 and has been used in the PDD field since that time (DoDPI, August 1992).

Defined as a control question test (CQT) the ZCT questions and pretesting procedures are designed to focus the examinee's "psychological set." The theory behind psychological set is that an examinee will focus his/her attention on the test question which holds the greatest significance, or signal value. Though there is considerable controversy concerning the accuracy of the CQT, it is the most widespread PDD examination technique in use today (OTA, 1983).

There are five types of questions contained in the ZCT: (a) relevant, (b) sacrifice relevant, (c) control, (d) irrelevant, and (e) symptomatic (DoDPI, August, 1992). Relevant questions deal with the specific issue or offense under investigation. According to CQT doctrine, the relevant questions pose the greatest threat to the guilty individual and will therefore produce the reactions having the greatest magnitude during a PDD examination. Relevant questions are further broken down into the categories of: (a) strong, (b) weak, and (c) evidence connecting.

The ZCT format uses two strong relevant questions (e.g., "Did you steal any of that money?", and "Did you steal any of that money from the Country Store?") and one evidence connecting question (e.g., "Do you know where any of that stolen money is now?").

Also included in the ZCT is a sacrifice relevant question. It is used to introduce the offense being probed but is not normally evaluated. Wording of the sacrifice relevant question is similar to the relevant question (e.g., "Regarding that stolen money, do you intend to answer truthfully each question about that?").

Control questions address an issue parallel to, but unrelated to the offense under investigation. For example, a suspect in a theft would be asked about prior theft activity. His physiological responses to the control questions would then be compared to his responses on the relevant questions. According to CQT doctrine, the signal value of the relevant questions will produce responses of a greater magnitude in guilty individuals, while the innocent individual--free of contrition for the crime--will instead focus attention on the control questions. There are two types of CQTs: (a) probable lie control (PLC), and (b) directed lie control (DLC).

A PLC question, the most widely used of the two techniques, presumes that all persons have engaged in a minor transgression similar to the offense under investigation. The PDD examiner capitalizes on that notion and essentially states that any individual capable of committing the minor transgression is also capable of committing the more serious crime. In effect, the innocent examinee is likely to be lying (hence the term probable lie) when he/she denies engaging in the behavior suggested in the control questions, while being truthful when denying participation in the relevant criminal behavior.

Use of the DLC requires that the examinee admit to committing some minor transgression, but is then instructed to lie about it during the PDD examination. In some instances, the same control question can be used for both techniques (e.g., "Prior to 1993, did you ever steal anything?"). The major difference between the techniques is in the handling during the pretest interview phase of the exam when the controls are being "set". Critics argue the benefits of one technique while vehemently denouncing the other.

Irrelevant questions are inconsequential to the criminal issue being investigated. They are typically used near the beginning of a PDD test to allay the examinee's reaction, or orienting response, at suddenly having the test begin. Additionally, they can be inserted throughout an examination when a substantial reaction or distortion has prevented the physiological tracings from returning to baseline levels. Unlike

the relevant and control questions, the irrelevant question is usually worded so that the examinee will answer yes, (e.g., "Are the lights on in this room?").

The final type of question found on the ZCT is referred to as a symptomatic question. It is designed to probe for an outside issue that could be more significant for an examinee than that for which he/she is being tested. When the symptomatic questions are briefed during the pretest interview, they are normally linked to the issue of examiner/examinee trust, (e.g., "Are you completely convinced that I will not ask you a question on this test that has not already been reviewed?", and "Is there something else you are afraid I will ask you a question about even though I have told you I would not?").

In summary, PASS was developed and validated using PDD field specific measures such as the ZCT, and has as a basis the standard scoring criteria both taught at DoDPI and used throughout the PDD community. In addition, the algorithm has been judged by APL to be accurate (92.9%-100.0% depending upon the measure) in distinguishing DI and NDI PDD examinations. The research described in this report was conducted to evaluate the effectiveness of the PASS when scoring controlled laboratory, or mock crime data where ground truth is known. Comparisons were also made between PASS accuracy and that of the PDD examiners who conducted the examinations.

### Method

This research was conducted onsite at the Department of Defense Polygraph Institute (DoDPI), Fort McClellan, AL. Data collection required 24 days during a consecutive five-week period. Personnel involved in the data collection process included two psychophysiological detection of deception (PDD) examiners, one subject handler, one role player known as the "deliberate intruder," and one scenario setter. The procedures employed are discussed below.

### Research Design

This research used a Zone Comparison Test (ZCT) format, and employed a 2 X 2 design as depicted in Figure 2. Subjects were programmed as innocent or guilty in a mock crime scenario involving the theft of \$124.00. Two types of control question tests (CQTs) were used during the study: (a) an experimental version of the directed lie control (DLC), and (b) the conventional probable lie control (PLC) currently in use throughout the PDD community.

The design was counterbalanced to prevent order effects. Examiner One began the project utilizing the DLC while Examiner Two began the project utilizing the PLC. Upon completing the data cell (approximately halfway through their respective data collection taskings) the PDD examiners switched question techniques (see Figure 3). Data from the first four subjects collected following the switch in question techniques was designated as piloting data and eliminated from the analysis.

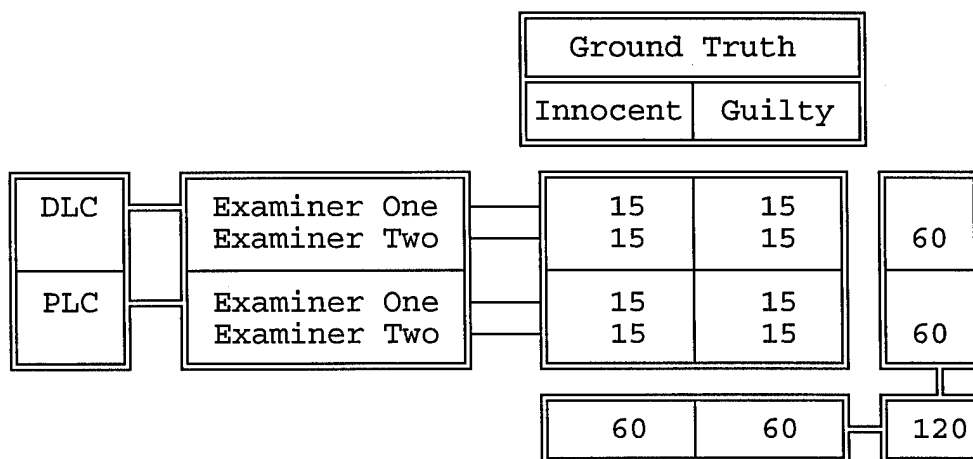


Figure 2. Diagram showing experimental design. DLC = directed lie control--experimental version; PLC = probable lie control.

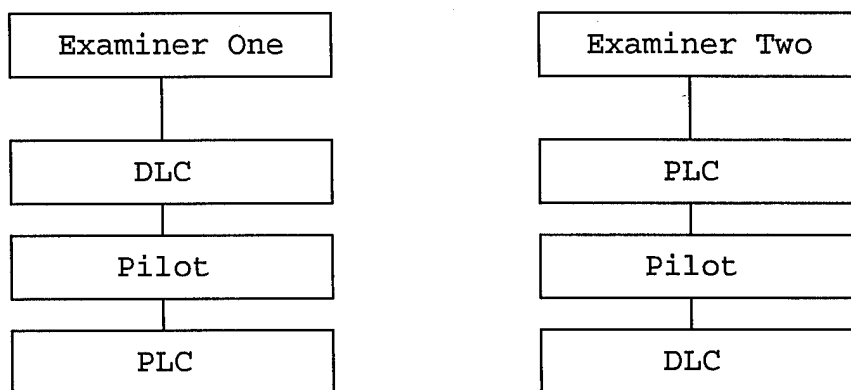


Figure 3. Diagram showing examiner data collection sequence. DLC = directed lie control--experimental version; PLC = probable lie control.

### Subjects

A total of 144 civilian citizens from the local community served as subjects during data collection. They were provided by an employment agency contracted for the recruitment of subjects, and were required to have had no prior PDD examinations. The data of 120 subjects were used in the analysis cited in this report. As illustrated in Table 5, the predominantly white (63.9%) group ranged in age from 19 to 60, and the majority of them (78.8%) had up to a high school education.

Based on self report the individuals were in good to excellent health (93.3%), and the majority were well rested, having had six or more hours of sleep the previous night (78.4%). As a group they were experiencing little pain or discomfort (98.4%), and a relatively small percentage (9.2%) indicated the use of medications prior to the examination (see Table 6).

Rationale for the elimination of 24 subjects from the analysis is detailed in this paragraph and summarized in Table 7. During mid-study the PDD examiners were required by the research design to change control question techniques (i.e., DLC to PLC, and vice versa). This change warranted a piloting period to become comfortable with administering the new technique. Each examiner conducted four mid-study piloting exams. Those exams were eliminated from the analyses. In addition, two individuals were unable to grasp the directed lie concept and on numerous occasions failed to correctly answer the control question during the examination. One set of test data was eliminated when the individual showed signs of a concussion as the result of a fall prior to reporting for the exam. A single incident of pretest confession occurred during the study and was eliminated. The last two exams conducted were also eliminated, having been designated to replace unusable data collected earlier in the day, if necessary. An additional ten examinations were eliminated due to poor electrodermal tracings, i.e., an internal scoring threshold established by the Polygraph Automated Scoring System (PASS) prevented PASS from being able to provide a score.

### PDD Examiners

All of the examinations were conducted by two certified PDD examiners assigned to DoDPI (see Table 8). The participating PDD examiners were proficient in operational procedures of the Axciton Computerized Polygraph equipment used during data collection. After three days of procedural refinements, each of the two PDD examiners conducted three exams per day until he completed his required number of subjects. The research design required that each examiner test 30 individuals who had been randomly programmed innocent, and 30 who had been randomly programmed guilty.

Table 5  
Subject Demographics (N = 120)

Category	Percent
Sex	
Female	57.5
Male	42.5
Age	
19-24	48.3
25-30	20.8
31-36	17.5
37-42	10.0
43-48	2.5
49-54	0.0
55-60	0.8
Race	
Asian	0.8
Black	32.8
Hispanic	2.5
White	63.9
Education	
No Diploma	24.6
High School	54.2
Vocational	5.1
Undergraduate	11.9
Graduate	4.2

Table 6  
Physical Condition of Subjects (N = 120)

Category	Percent
Health	
Excellent	45.0
Good	48.3
Fair	6.7
Medication	
No	90.8
Yes	9.2
Discomfort	
None	94.2
Mild	4.2
Moderate	1.7
Sleep (hrs)	
0.0 - 2.9	3.3
3.0 - 5.9	18.3
6.0 - 8.9	69.2
9.0 - 11.9	9.2

Table 7  
Data Excluded From Analysis

Rationale	# Exams
Mid-study piloting	8
Confused by test format	2
Medical elimination	1
Pretest confession	1
Unused backup data	2
Poor electrodermal	10
Total	24



Table 8  
PDD Examiner Qualifications

Assignment	Years Experience	
	Examiner One	Examiner Two
Criminal Investigator	9	11
PDD Examiner (field)	3	6
DoDPI Instructor	6	2
Total	18	19

Note. PDD = psychophysiological detection of deception;  
DoDPI = Department of Defense Polygraph Institute.

#### Apparatus

Hardware. Two Axciton Computerized Polygraph Systems (Version 48-I; 16 bit parallel format) were used. The specific channels consisted of: (a) two pneumograph channels utilizing convoluted tubes to measure changes in thoracic and abdominal areas during expiration and inspiration, (b) one electrodermal channel utilizing fingerplate electrodes to measure changes in sweat gland activity on two fingers of the subject's non-dominant hand, and (c) one cardiograph channel utilizing a standard medical blood pressure cuff, pump bulb assembly and sphygmomanometer to indicate changes in relative blood pressure.

All examinations conducted during this project were administered in standard configuration PDD suites maintained by DoDPI. Each suite contained an examiner desk with an Axciton Computerized Polygraph system, a two-way mirror, wall and ceiling mounted video cameras, and a chair for both the examiner and examinee, along with a wide-armed, high-backed examination chair. All sessions were videotaped using wall and ceiling mounted video cameras and commercial videotape recorders.

Software. The Polygraph Automated Scoring System (PASS), Version 2.0 developed by the Johns Hopkins University Applied Physics Laboratory (APL), (1993b) was used to analyze the physiological data collected and stored by the Axciton Computerized Polygraph.

Crunch, Version 4.0, Crunch Software Corporation (1991), was used to accomplish the statistical analysis of the data. The FREQ, TABLES and CROSSTAB modules were used to generate percentages, chi-square and Fisher exact probability calculations.

### Crime Scene

Space within a typical office lounge was used as the crime scene. The area was referred to as the Country Store and consisted of two small tables, one of which contained a display of snack food (i.e., cookies, candy, etc.). The other table held a small plastic box which contained \$124.00 in paper currency (four \$1; two \$5; one \$10; and one \$100), and \$3.00 in assorted coins. The paper currency was stacked in a specific order and configuration before being placed in the box to make it difficult to be counted quickly.

### Scenario

The mock crime was defined as a theft of \$124.00 from the Country Store cash box. To prevent the PDD examiners from discerning guilt or innocence via the individual's knowledge of the building, both innocent and guilty subjects traveled the same route from briefing area to exam room.

Innocent Subjects. The scenario setter took each subject from the waiting area individually, escorted him/her to the designated briefing area, and utilized the script in Appendix B to accomplish the programming. All were told that a theft of money from the Country Store had occurred, and that each were suspects in the case due to having been in the area at the time. Each subject was assured that he/she was innocent of the crime and that the task at hand was simply to be honest and cooperative with the PDD examiner.

Following programming, each subject was escorted to a holding area with a stop along the way to pick up score sheets from a forms closet. As soon as possible following programming each subject was escorted to the PDD examination room and introduced to the PDD examiner.

Guilty Subjects. The scenario setter took each subject from the waiting area individually, escorted him/her to the designated briefing area, and used the script in Appendix C to accomplish the programming. In turn, all were told that they were going to steal money from the Country Store and then undergo a PDD examination regarding the theft of the money. It was explained to each subject that the primary goal was to convince the PDD examiner that he/she was innocent of committing the crime. Each individual was then escorted into the area referred to as the Country Store, and shown the cash box and store merchandise.

The scenario setter explained and demonstrated the steps each subject was to follow in committing the theft: (a) take only the paper currency out of the box, (b) count it, (c) conceal it on their person or in a purse, and then (d) immediately leave the room through the designated doorway. It was stressed to all guilty participants that it was vitally important not to be seen stealing the money, otherwise the PDD examiner could inadvertently be informed of the circumstances prior to the exam,

effectively rendering the exam unnecessary. (NOTE: All subjects had been informed that the employment agency would not pay them if they did not take the PDD exam.)

If seen with the money, each subject was instructed to act as if he/she had just purchased a candy bar and were simply making change. Each individual was then to conceal the money and leave the room immediately with a candy bar in hand. Following a review of the steps, each subject was then left alone in the room to carry out the scenario.

Approximately 10-15 seconds after the scenario setter left the room an individual known as the deliberate intruder entered the crime scene through another doorway in order to surprise the subject committing the theft. The deliberate intruder was instructed to remain in the room, making small talk, cleaning the sink counter, etc., until the individual completed the task and left. This was done to heighten the arousal level of the subject by requiring him/her to conceal his actions while committing the crime. The deliberate intruder was not there to confront the individual, but rather to make him/her nervous by obstructing a "clean" getaway following the commission of the crime.

When the subject left the room and joined the scenario setter he/she was lead to a forms closet, where the scenario setter made sure the individual both had the money, and knew the amount of money stolen. If the subject had not completed counting the money prior to being interrupted he/she was instructed to confirm the amount at that time. The money was again concealed on the subject's person, and the candy bar was hidden in a box to prevent the examiner from realizing that the individual had been in the Country Store.

In order to reinforce details regarding the commission of the theft, the subject was then escorted to a holding area where he/she completed the "Crime Scene" Scenario questionnaire in Appendix D. As soon as possible following programming each subject was escorted to the PDD examination room and introduced to the PDD examiner.

#### Procedure

The principal steps in the daily data collection process are shown in Figure 4. Upon arrival, the subjects were welcomed to DoDPI and provided with a general briefing on the purpose of the study (see Appendix E). Following the briefing, the subjects were asked about prior PDD experience. Those who had prior PDD experience were not allowed to participate in this project.

Subjects were then provided with a packet containing a copy of the Project Briefing (see Appendix F), and the Volunteer Agreement Affidavit (see Appendix G). Both documents informed the subject that his/her participation was completely voluntary. After signing the Volunteer Agreement Affidavit the subjects

completed the Background Information Form (see Appendix H) also included in the packet. Subjects who chose to participate in the study were then escorted to a designated waiting area.

Each subject was randomly assigned to either the innocent group or the guilty group and was then programmed individually as described in the Scenario section of this report. (NOTE: Programmed subjects were kept separated from all other programmed and non-programmed subjects until the PDD examination had been completed.) Subjects were then introduced to a PDD examiner, who administered the pretest and in-test portions of a ZCT PDD examination.

As stated previously, two question techniques were used in this project: the DLC and the PLC. The philosophies behind the two techniques required that the interview conducted during the pretest phase be structured differently. Appendices I and J, respectively, contain the general framework for the pretest interview for the two question techniques. Also occurring during the pretest phase was the review and signing of both a Rights Advisement Form (see Appendix K), and a Consent to Interview With Polygraph form (see Appendix L).

The in-test phase required the use of the Stimulation (numbers) Test (see Appendix M), and the ZCT Question List for a mock "theft" scenario (see Appendix N). The PDD examiner administered the numbers test and then briefly discussed it with the subject in order to focus attention on the individual's physiological changes which occurred during the commission of the lie.

Utilizing the ZCT question list, the PDD examiner collected three tests, with a fourth test authorized only if necessary, (i.e., as a result of movement distortions, detected countermeasures, etc.). Also, DoDPI (August 1992) guidelines for administering the ZCT allowed the PDD examiners to rotate the control questions between tests in an attempt to place the strongest reacting control question in position #6.

The PDD examiner was not allowed to talk to or prompt the subject between tests. If asked a question by the subject, the PDD examiner attempted to answer the question in such a way as not to jeopardize the outcome of the project. Following the PDD examination, each subject was debriefed (see Appendix O), and released.

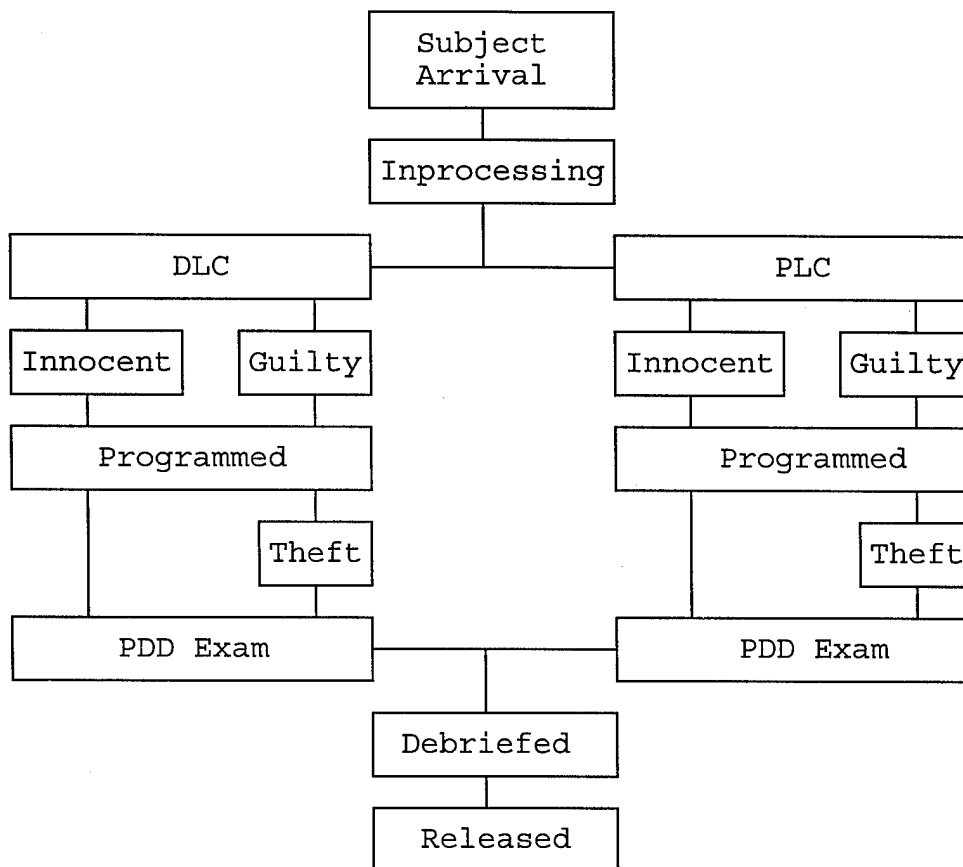


Figure 4. Flow chart of principal steps in data collection process. DLC = directed lie control--experimental version; PLC = probable lie control; PDD = psychophysiological detection of deception.

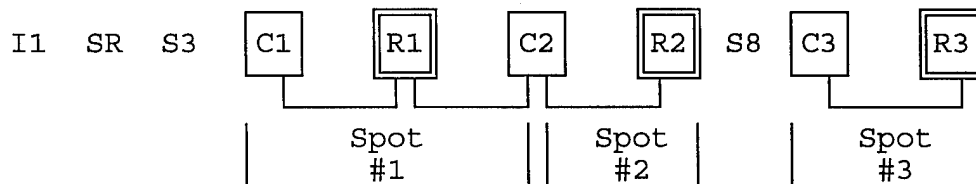
#### Scoring Criteria and Data Analysis Procedures

PDD Examiner. Each examination was scored by the original examiner, and later blind scored by two other certified and similarly qualified PDD examiners. Both original examiners scored their charts either immediately following each examination or at the end of the day. A standard seven position scale (+3, +2, +1, 0, -1, -2, -3) was used in conjunction with the DoDPI (August 1992) criteria for ZCT spot analysis, and numerical evaluation (see Figures 5 and 6, respectively).

PASS. As mentioned earlier the PASS cut score during the development phase was 80-20. However, prior to the distribution of the PASS Version 2.0 prototype software, APL changed the cut score to 90-10. Therefore, any examination during this study which received a probability score of 0.90 or higher was

PASS is designed to recognize, tag, and eliminate segments within the physiological tracings which it defines as artifacts. However, PASS enables, and in fact occasionally requires subjective manipulation of the data in the area of artifact editing. The examiner can override PASS's decision to eliminate an area from scoring consideration, and can also select areas to be eliminated which PASS did not tag. As a result, each data file in this study was scored twice.

On the first run, PASS was allowed to interpret the data without benefit of subjective manipulation, i.e., tagging only those areas which violated its definition of acceptable. On the second run, problem areas identified by the original examiners and the researcher were edited, or tagged for elimination. On the few occasions (17) which resulted in a split decision the call from the second run (reflecting examiner edits) was used.



**Figure 5.** Diagram showing Zone Comparison Test (ZCT) question sequence and evaluation spots with compared control and relevant questions linked. I = irrelevant; SR = sacrifice relevant; S = symptomatic; C = control, and; R = relevant.

Spot#		Score	Call
1 + 2 + 3	≥	+6 (with no spot equal to 0)	NDI
1 + 2 + 3	≤	-6 (with no spot equal to 0)	DI
any	=	-3	DI
1 + 2 + 3	=	any score not mentioned above	INC

Figure 6. Chart showing numerical evaluation criteria for Zone Comparison Test (ZCT). NDI = no deception indicated; DI = deception indicated; INC = inconclusive.

The objective of this project was to determine the effectiveness of the Polygraph Automated Scoring System (PASS),

Version 2.0 prototype software in detecting deception in a controlled laboratory study using mock crime data. For the purposes of comparison, all accuracy findings cited in this report were structured to parallel the type of assessments made by the Applied Physics Laboratory (APL) during algorithm development, i.e., accuracy rates for the data set overall, along with a recalculated rate excluding the inconclusive (INC) decisions, as would be done in field data reporting.

PASS accuracy, and PDD examiner accuracy for this study were compared to ground truth. A chi-square test was performed for the data citing decisions of NDI, DI and INC. A Fisher exact probability (2-tail) test was performed for the data when INC decisions were eliminated. Both chi-square and Fisher exact probability statistics are summarized following each section of accuracy comparisons.

In addition, analyses regarding specific issues germane to the results generated by this research were examined. Those issues include: (a) differences between the types of CQTs used, (b) interrater reliability, and (c) analysis of PDD examiner accuracy over time. PASS and PDD examiner agreement was also assessed to enable a direct comparison with the data generated during PASS development and validation. Data from 120 subjects were used in the following analyses.

#### Overall Accuracy Comparisons

PASS Decision vs. Ground Truth. The results shown in Table 9 indicate that across all conditions for the 120 cases PASS accurately identified 62.5% of the subjects as programmed. The proportion of cases for which PASS generated an incorrect, or reverse decision, was 16.7%, and the INC rate was 20.8%. Table 9 reflects that on the cases where PASS was able to make a definitive call (i.e., INCs eliminated), the rate of agreement between PASS and ground truth was 79.0% and the percentage of incorrect decisions was to 21.1%.

Table 9  
Percentage of Agreement Between PASS Decision and Ground Truth--All Cases (N = 120)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Agreed	62.5	(75)	79.0	(75)
Disagreed (FN + FP)	16.7	(20)	21.1	(20)
Inconclusive	20.8	(25)	--	--

Note. FN = false negative; FP = false positive; PASS = Polygraph Automated Scoring System.

PDD Examiner Decision vs. Ground Truth. The overall scoring accuracy of the two PDD examiners who collected the data was 71.1% (see Table 10). The examiners arrived at a reverse decision in 18.3% of the cases, and had an INC rate of 10.0%. This gave the PDD examiners a higher overall rate of agreement and a lower percentage of INC calls than PASS. Again eliminating the INC decisions from the calculations, the rate of PDD examiner agreement with ground truth was adjusted to 79.6%, with incorrect decisions totaling 20.4%; essentially equivalent to PASS's performance under the same circumstances. NOTE: Chi-square ( $X^2$ ) and Fisher exact test results for comparisons of accuracy between PDD examiners were not significant, when computed with and without INC cases, respectively,  $X^2(2, N = 120) = .057, p = 0.9719$  and Fisher exact test ( $N = 108$ ),  $p = .8116$ . Thus indicating that the decisions made by the two PDD examiners were not associated.

Table 10

Percentage of Agreement Between PDD Examiner Decision and Ground Truth--All Cases

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Agreed	71.7	(86)	79.6	(86)
Disagreed (FN + FP)	18.3	(22)	20.4	(22)
Inconclusive	10.0	(12)	--	--

Note. FN = false negative; FP = false positive;  
PDD = psychophysiological detection of deception.

Chi-Square and Fisher Exact Statistics. A chi-square test was used to assess the proportional differences between correct, incorrect and INC decisions made by PASS and the PDD examiners. Table 11 indicates that the differences between the accuracy of PASS and PDD examiner decisions in this study were not statistically significant, i.e., the accuracy of the decisions made by PASS was independent of the accuracy of the decisions made by the PDD examiner.

The INC decisions were then eliminated and a Fisher exact probability test was applied to only the correct and incorrect decisions made by PASS and the PDD examiners. The results of the Fisher exact test were also not significant.



Table 11

Results of Chi-square ( $X^2$ ) and Fisher Exact Statistics when Testing Association Between PASS and PDD Examiner Decisions-- All Cases

Comparison	df	n	Value	p	Significant ( $\alpha = .05$ )
PASS Decision vs. PDD Examiner Decision					
$X^2$ <sup>a</sup>	2	240	4.537	.1035	No
Fisher Exact <sup>b</sup>		203		.9999	No

Note. PASS = Polygraph Automated Scoring System; PDD = psychophysiological detection of deception.

<sup>a</sup> Test corrected for continuity. <sup>b</sup> Probability test (2-tail).

NDI and DI Accuracy Comparisons

PASS Decision vs. Ground Truth. The results shown in Table 12 demonstrate that PASS was more accurate in clearing programmed innocent individuals, rather than detecting programmed guilty individuals. PASS accurately identified 71.7% of the 60 innocent individuals and 53.3% of the 60 guilty individuals. The corresponding false positive (FP) and false negative (FN) rates were 10.0% and 23.3%, respectively. PASS had an INC rate of 18.3% on innocent examinees and 23.3% on guilty examinees.

PDD Examiner Decision vs. Ground Truth. Converse to the results cited for PASS, Table 13 shows that the PDD examiners were more accurate in detecting programmed guilty individuals, than in clearing programmed innocent individuals. The PDD examiners also had a much lower INC rate for the innocent cases than PASS (5.0% compared to 23.3%). The PDD examiners accurately identified 61.7% of the 60 individuals programmed innocent and 81.7% of the 60 individuals programmed guilty. The corresponding FP and FN rates were 23.3% and 13.3%, respectively. The PDD examiners had an INC rate of 15.0% on innocent examinees and 5.0% on guilty examinees.

Table 12

Percentage of Agreement Between PASS Decision and Ground Truth--Programmed Innocent and Programmed Guilty Cases  
(N = 120)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Innocent ( <u>n</u> = 60)				
Agreed	71.7	(43)	87.8	(43)
Disagreed (FP)	10.0	(6)	12.3	(6)
Inconclusive	18.3	(11)	--	--
Guilty ( <u>n</u> = 60)				
Agreed	53.3	(32)	69.6	(32)
Disagreed (FN)	23.3	(14)	30.4	(14)
Inconclusive	23.3	(14)	--	--

Note. FN = false negative; FP = false positive; PASS = Polygraph Automated Scoring System.

Table 13

Percentage of Agreement Between PDD Examiner Decision and Ground Truth--Programmed Innocent and Programmed Guilty Cases (N = 120)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Innocent ( <u>n</u> = 60)				
Agreed	61.7	(37)	72.6	(37)
Disagreed (FP)	23.3	(14)	27.5	(14)
Inconclusive	15.0	(9)	--	--
Guilty ( <u>n</u> = 60)				
Agreed	81.7	(49)	86.0	(49)
Disagreed (FN)	13.3	(8)	14.0	(8)
Inconclusive	5.0	(3)	--	--

Note. FN = false negative; FP = false positive;  
PDD = psychophysiological detection of deception.

Chi-Square and Fisher Exact Statistics. Table 14 shows chi-square and Fisher exact probability test results when accuracy on programmed innocent cases versus programmed guilty cases was assessed individually for both PASS and the PDD examiners. Using correct, incorrect and INC decisions to compute chi-square, the

result indicated that the accuracy of the decision rendered by PASS was independent of the programmed status of the individual. The same statement is true for the PDD examiners.

With the INC decisions eliminated, the Fisher exact probability test was statistically significant ( $p = .0432$ ) in regard to PASS's rate of accuracy on innocent versus guilty subjects. Conversely, the results showed that the accuracy of the decisions produced by the PDD examiners were independent of the programmed condition of the individual.

Table 14

Results of Chi-square ( $X^2$ ) and Fisher Exact Statistics when Testing Association Between PASS and PDD Examiner Decisions--Programmed Innocent and Programmed Guilty Cases

Comparison	df	n	Value	p	Significant ( $\alpha = .05$ )
PASS Decision--Innocent vs. Guilty					
$X^2$ <sup>a</sup>	2	120	3.943	.1392	No
Fisher Exact <sup>b</sup>		95		.0432	Yes
PDD Examiner Decision--Innocent vs. Guilty					
$X^2$ <sup>a</sup>	2	120	4.627	.0989	No
Fisher Exact <sup>b</sup>		108		.0982	No
PASS vs. PDD Examiner--Innocent Cases					
$X^2$ <sup>a</sup>	2	120	2.813	.2451	No
Fisher Exact <sup>b</sup>		100		.0799	No
PASS vs. PDD Examiner--Guilty Cases					
$X^2$ <sup>a</sup>	2	120	10.179	.0062	Yes
Fisher Exact <sup>b</sup>		103		.0546	No

Note. PASS = Polygraph Automated Scoring System; PDD = psychophysiological detection of deception.

<sup>a</sup> Test corrected for continuity. <sup>b</sup> Probability test (2-tail).

#### Control Question Test (CQT) Comparisons

PASS Decision vs. Directed Lie Control (DLC). In order to detect any bias created by the two CQTs, analyses were conducted to assess differences between PASS performance on the DLC and PLC. Table 15 shows that for the data collected using the experimental version of the DLC, PASS had an overall accuracy of 61.7%. The percentage of incorrect decisions using the DLC was 15.0% and the INC rate was 23.3%. Elimination of the high number of INC calls resulted in a dramatic upward adjustment (80.4%) in the rate of agreement between PASS and ground truth.

When the DLC cases were broken down by innocent and guilty programming Table 16 shows that PASS was more accurate at identifying the innocent subjects than the guilty subjects (66.7% and 56.7%, respectively). There was also a higher INC rate on the guilty subjects. When INC decisions were eliminated there

was a resulting accuracy assessment of 83.3% for innocent subjects and 77.3% for guilty individuals.

Table 15

Percentage of Agreement Between PASS Decision and Ground Truth--All Directed Lie Control (DLC) Cases (N = 60)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Agreed	61.7	(37)	80.4	(37)
Disagreed (FN + FP)	15.0	(9)	19.6	(9)
Inconclusive	23.3	(14)	--	--

Note. FN = false negative; FP = false positive; PASS = Polygraph Automated Scoring System.

Table 16

Percentage of Agreement Between PASS Decision and Ground Truth--Programmed Innocent and Programmed Guilty Cases using Directed Lie Control (DLC) (N = 60)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Innocent ( <u>n</u> = 30)				
Agreed	66.7	(20)	83.3	(20)
Disagreed (FP)	13.3	(4)	16.7	(4)
Inconclusive	20.0	(6)	--	--
Guilty ( <u>n</u> = 30)				
Agreed	56.7	(17)	77.3	(17)
Disagreed (FN)	16.7	(5)	22.7	(5)
Inconclusive	26.7	(8)	--	--

Note. FN = false negative; FP = false positive; PASS = Polygraph Automated Scoring System.

PASS Decision vs. Probable Lie Control (PLC). PASS's overall accuracy on the PLC cases was comparable to that on the DLC cases, though the PLC INC rate was slightly lower (18.3% compared to 23.3%) and the rates of agreement and disagreement were slightly higher (63.3% compared to 61.7%, and 18.3% compared to 15.0%, respectively) (see Table 17).

The innocent/guilty breakout analyses shown in Table 18 reveal a disparity between PASS's accuracy for identifying innocent (76.7%) versus guilty (50.0%) individuals. The FN rate

for programmed guilty cases was accordingly higher (30.0% compared to 6.7%), as was the INC rate (20.0% compared to 16.7%).

Chi-Square and Fisher Exact Statistics. Table 19 shows chi-square test results for PASS accuracy on the DLC and PLC cases. First, the analyses compared PASS's overall accuracy on the DLC versus the PLC. Additionally, PASS accuracy on the innocent and the guilty cases was compared both between CQTs and within each CQT. None of the chi-square tests were significant, indicating that all compared variables were independent. However, Fisher exact probability test results for the same comparisons were statistically significant when examining PASS's accuracy on the innocent and guilty cases when using the PLC.

PDD Examiner Decision vs. DLC. Comparable analyses were performed for PDD examiner calls to detect any bias created by the two CQTs. Table 20 shows that the PDD examiners had an overall agreement rate with ground truth in 70.0% of the cases. They arrived at incorrect decisions 15.0% of the time and also had an INC rate of 15.0%.

When comparing the breakout data for innocent and guilty individuals, Table 21 shows that the DLC accuracy rate for PDD examiners on innocent cases was just 50.0%, with FN decisions accounting for 23.3% and INC calls totaling 26.7%.

PDD examiner performance on the DLC guilty cases was markedly different from the innocent cases. The accuracy rate for detecting individuals who had committed the theft was 90.0%. The FP rate was 6.7% and the INC rate was 3.3%. (NOTE: These figures are inflated due to the small n used in calculation.)

Table 17  
Percentage of Agreement Between PASS Decision and Ground Truth--All Probable Lie Control (PLC) Cases (N = 60)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Agreed	63.3	(38)	77.6	(38)
Disagreed (FN + FP)	18.3	(11)	22.5	(11)
Inconclusive	18.3	(11)	--	--

Note. FN = false negative; FP = false positive; PASS = Polygraph Automated Scoring System.

Table 18

Percentage of Agreement Between PASS Decision and Ground Truth--Programmed Innocent and Programmed Guilty Cases using Probable Lie Control (PLC) (N = 60)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Innocent ( <u>n</u> = 30)				
Agreed	76.7	(23)	92.0	(23)
Disagreed (FP)	6.7	(2)	8.0	(2)
Inconclusive	16.7	(5)	--	--
Guilty ( <u>n</u> = 30)				
Agreed	50.0	(15)	62.5	(15)
Disagreed (FN)	30.0	(9)	37.5	(9)
Inconclusive	20.0	(6)	--	--

Note. FN = false negative; FP = false positive; PASS = Polygraph Automated Scoring System.

Table 19

Results of Chi-square ( $X^2$ ) and Fisher Exact Statistics when Testing Association Between Directed Lie Control (DLC) and Probable Lie Control (PLC) Cases--PASS Decisions

Comparison	df	<u>n</u>	Value	p	Significant ( $\alpha = .05$ )
DLC Format vs. PLC Format					
$X^2$ <sup>a</sup>	2	120	0.210	.9003	No
Fisher Exact <sup>b</sup>		95		.8043	No
DLC vs. PLC - Innocent Cases					
$X^2$ <sup>a</sup>	2	60	0.260	.8782	No <sup>c</sup>
Fisher Exact <sup>b</sup>		49		.4174	No
DLC vs. PLC - Guilty Cases					
$X^2$ <sup>a</sup>	2	60	0.746	.6888	No
Fisher Exact <sup>b</sup>		46		.3457	No
DLC Format - Innocent vs. Guilty					
$X^2$ <sup>a</sup>	2	60	0.180	.9141	No
Fisher Exact <sup>b</sup>		46		.7178	No
PLC Format - Innocent vs. Guilty					
$X^2$ <sup>a</sup>	2	60	4.562	.1022	No
Fisher Exact <sup>b</sup>		49		.0181	Yes

Note. PASS = Polygraph Automated Scoring System.

<sup>a</sup> Test corrected for continuity. <sup>b</sup> Probability test (2-tail). <sup>c</sup> Minimum expected frequency was 3.0. Two cells had expected frequencies of less than 5, therefore chi-square p-values may not be accurate.

Table 20

Percentage of Agreement Between PDD Examiner Decision and Ground Truth--All Directed Lie Control (DLC) Cases (N = 60)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Agreed	70.0	(42)	82.4	(42)
Disagreed (FN + FP)	15.0	(9)	17.7	(9)
Inconclusive	15.0	(9)	--	--

Note. FN = false negative; FP = false positive; PDD = psychophysiological detection of deception.

Table 21

Percentage of Agreement Between PDD Examiner Decision and Ground Truth--Programmed Innocent and Programmed Guilty Cases using Directed Lie Control (DLC) (N = 60)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Innocent ( <u>n</u> = 30)				
Agreed	50.0	(15)	68.2	(15)
Disagreed (FP)	23.3	(7)	31.8	(7)
Inconclusive	26.7	(8)	--	--
Guilty ( <u>n</u> = 30)				
Agreed	90.0	(27)	93.1	(27)
Disagreed (FN)	6.7	(2)	6.9	(2)
Inconclusive	3.3	(1)	--	--

Note. FN = false negative; FP = false positive; PDD = psychophysiological detection of deception.

PDD Examiner Decision vs. PLC. Table 22 shows that both PDD examiner agreement (73.3% versus 70.0%) and disagreement (21.7% versus 15.0%) were higher using the PLC technique than when using the DLC technique. The PDD examiner INC rate was also lower on the PLC than on the DLC (5.0% compared to 15.0%).

PDD examiner accuracy on the PLC was essentially equivalent for both innocent and guilty cases (see Table 23). They had an agreement rate of 73.3% on both the programmed innocent and the programmed guilty individuals. The incorrect decisions were slightly higher on the innocent cases than on the guilty cases (23.3% compared to 20.0%), however the INC rate was lower on the

innocent cases than the guilty cases (3.3% compared to 6.7%)  
(NOTE: These figures are inflated due to the small n used in calculation.)

Table 22

Percentage of Agreement Between PDD Examiner Decision and Ground Truth--All Probable Lie Control (PLC) Cases (N = 60)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Agreed	73.3	(44)	77.2	(44)
Disagreed (FN + FP)	21.7	(13)	22.8	(13)
Inconclusive	5.0	(3)	--	--

Note. FN = false negative; FP = false positive; PDD = psychophysiological detection of deception.

Table 23

Percentage of Agreement Between PDD Examiner Decision and Ground Truth--Programmed Innocent and Programmed Guilty Cases using Probable Lie Control (PLC) (N = 60)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Innocent ( <u>n</u> = 30)				
Agreed	73.3	(22)	75.9	(22)
Disagreed (FP)	23.3	(7)	24.1	(7)
Inconclusive	3.3	(1)	--	--
Guilty ( <u>n</u> = 30)				
Agreed	73.3	(22)	78.6	(22)
Disagreed (FN)	20.0	(6)	21.4	(6)
Inconclusive	6.7	(2)	--	--

Note. FN = false negative; FP = false positive; PDD = psychophysiological detection of deception.

Chi-Square and Fisher Exact Statistics. The chi-square results shown in Table 24 reveal that PDD examiner accuracy was independent of the CQT used, as well as the programmed status of the individual. The results were, however, statistically significant ( $p = .0132$ ) for PDD examiner accuracy on innocent and guilty individuals when using the DLC. The same was true for the Fisher exact test when the INC decisions were eliminated ( $p = .0293$ ). The remaining Fisher exact tests results were not significant.



Table 24

Results of Chi-square ( $X^2$ ) and Fisher Exact Statistics when Testing Association Between Directed Lie Control (DLC) and Probable Lie Control (PLC) Cases--PDD Examiner Decisions

Comparison	df	n	Value	p	Significant ( $\alpha = .05$ )
DLC Format vs. PLC Format					
$X^2$ <sup>a</sup>	2	120	2.504	.2859	No
Fisher Exact <sup>b</sup>		108		.6336	No
DLC vs. PLC--Innocent Cases					
$X^2$ <sup>a</sup>	2	60	4.973	.0832	No <sup>c</sup>
Fisher Exact <sup>b</sup>		51		.7523	No
DLC vs. PLC--Guilty Cases					
$X^2$ <sup>a</sup>	2	60	1.452	.4840	No <sup>d</sup>
Fisher Exact <sup>b</sup>		57		.1443	No
DLC Format--Innocent vs. Guilty					
$X^2$ <sup>a</sup>	2	60	8.659	.0132	Yes
Fisher Exact <sup>b</sup>		51		.0293	Yes
PLC Format--Innocent vs. Guilty					
$X^2$ <sup>a</sup>	2	60	0.000	1.0000	No
Fisher Exact <sup>b</sup>		57		1.0000	No

Note. PDD = psychophysiological detection of deception.

<sup>a</sup> Test corrected for continuity. <sup>b</sup> Probability test (2-tail). <sup>c</sup> Minimum expected frequency was 4.5. Two cells had expected frequencies of less than 5 therefore chi-square p-values may not be accurate. <sup>d</sup> Minimum expected frequency was 1.5. Four cells had expected frequencies of less than 5 therefore chi-square p-values may not be accurate.

PASS/PDD Examiner Comparisons Using DLC/PLC Cases. In tables cited previously it was shown that the PDD examiners had a higher (70.0% compared to 61.7%) overall accuracy rate on the DLC than did PASS along with a lower INC rate (15.0% compared to 23.3%) (Table 15 and Table 20).

When comparing the breakout data for DLC innocent and guilty individuals, Table 16 and Table 21 showed that PASS had a higher rate of agreement on the innocent cases than did the PDD examiners (66.7% compared to 50.0%) and lower FP and INC rates (13.3% compared to 23.3%, and 20.0% compared to 26.7%, respectively). However, the PDD examiners far exceeded PASS's accuracy on the guilty cases (90.0% compared to 56.7%) and had lower FN and INC rates, as well (6.7% compared to 16.7%, and 3.3% compared to 26.7%, respectively).

Both PASS (Table 15 and Table 17) and PDD examiner (Table 20 and Table 22) rates of agreement were higher using the PLC technique versus the DLC technique (63.3% compared to 61.7%, and 73.3% compared to 70.0%, respectively).

The rate of PDD examiner agreement on the PLC was higher than that of PASS on the PLC (73.3% compared to 63.3%) (Table 17 and Table 22). The PDD examiner rate of incorrect decisions was also higher than PASS (21.7% compared to 18.3%), but the PDD examiner INC rate was lower (5.0% compared to 18.3%).

When comparing breakout data for the PLC innocent and guilty individuals Table 18 and Table 23 showed that PASS had a higher rate of agreement on the programmed innocent individuals than did the PDD examiners (76.7% compared to 73.3%). PASS also had a lower FP rate (6.7% compared to 23.3%), but the PDD examiners had a lower INC rate (3.3% compared to 16.7%). As with the case of the DLC comparison, the PDD examiners far exceeded PASS's accuracy on detecting programmed guilty individuals on the PLC (73.3% compared to 50.0%) and also had lower FN and INC rates than PASS (20.0% compared to 30.0%, and 6.7% compared to 20.0%, respectively).

Table 25

Results of Chi-square ( $X^2$ ) and Fisher Exact Statistics when Testing Association Between PASS and PDD Examiner--Programmed Innocent and Programmed Guilty Cases using Directed Lie Control (DLC)

Comparison	df	n	Value	p	Significant ( $\alpha = .05$ )
PASS vs. PDD Examiner--DLC--All Cases					
$X^2$ <sup>a</sup>	2	120	0.898	.6382	No
Fisher Exact <sup>b</sup>		97		1.0000	No
PASS vs. PDD Examiner--DLC--Innocent					
$X^2$ <sup>a</sup>	2	60	0.892	.6401	No
Fisher Exact <sup>b</sup>		46		.3068	No
PASS vs. PDD Examiner--DLC--Guilty					
$X^2$ <sup>a</sup>	2	60	6.412	.0405	Yes <sup>c</sup>
Fisher Exact <sup>b</sup>		51		.2163	No

Note. PASS = Polygraph Automated Scoring System; PDD = psychophysiological detection of deception.

<sup>a</sup> Test corrected for continuity. <sup>b</sup> Probability test (2-tail).

<sup>c</sup> Minimum expected frequency was 3.5. Four cells had expected frequencies of less than 5 therefore chi-square p-values may not be accurate.

Chi-Square/Fisher Exact Statistics. Table 25 and Table 26 show the chi-square and Fisher exact probability results computed for PASS and PDD examiner comparisons for both the DLC and PLC. Using chi-square there was an indication that decisions rendered by PASS and the PDD examiners were not independent when comparing the programmed guilty cases of the DLC, however that result is suspect due to smaller than expected frequencies in a number of

the cells. After eliminating the INC decisions from the analysis, the computed Fisher exact was not significant. All other comparisons between PASS and the PDD examiners were also not significant.

Table 26

Results of Chi-square ( $X^2$ ) and Fisher Exact Statistics when Testing Association Between PASS and PDD Examiner--Programmed Innocent and Programmed Guilty Cases using Probable Lie Control (PLC)

Comparison	df	n	Value	p	Significant ( $\alpha = .05$ )
PASS vs. PDD Examiner--PLC--All Cases					
$X^2$ <sup>a</sup>	2	120	3.847	.1461	No
Fisher Exact <sup>b</sup>		106		1.0000	No
PASS vs. PDD Examiner--PLC--Innocent					
$X^2$ <sup>a</sup>	2	60	3.278	.1942	No <sup>c</sup>
Fisher Exact <sup>b</sup>		54		.1531	No
PASS vs. PDD Examiner--PLC--Guilty					
$X^2$ <sup>a</sup>	2	60	2.365	.3066	No <sup>d</sup>
Fisher Exact <sup>b</sup>		52		.2337	No

Note. PASS = Polygraph Automated Scoring System; PDD = psychophysiological detection of deception.

<sup>a</sup> Test corrected for continuity. <sup>b</sup> Probability test (2-tail). <sup>c</sup> Minimum expected frequency was 3.0. Four cells had expected frequencies of less than 5, therefore chi-square p-values may not be accurate. <sup>d</sup> Minimum expected frequency was 4.2. Two cells had expected frequencies of less than 5, therefore chi-square p-values may not be accurate.

#### Interrater Agreement Analyses

Blind Scoring. Two PDD examiners with qualifications similar to the "original" examiners were asked to blind score the 120 examinations. They followed the same instructions and criteria for scoring as the PDD examiners who performed the data collection. However, they had no knowledge of the individual subjects associated with the exams, nor the decisions reached by the original examiner. The scorer referred to as Blind 1, had an accuracy rate of 60.83% when compared to ground truth. Blind 2 had an accuracy rate of 65.00%.

Kappa, a measure of interrater agreement was computed for the two blind scorers. Table 27 shows that the percentage of agreement between Blind 1 and Blind 2 was 69.17%. The percentage of agreement which could be expected by chance alone was 36.50%. Accordingly, the computed Kappa value indicated that the difference between Blind 1 and Blind 2 agreement was not due to chance ( $p = .0000$ ).

Kappa was also computed for each blind scorer when compared to the decisions generated by Examiner One, Examiner Two, and PASS. All values were statistically significant. The level of accuracy for Blind 1 and Blind 2 was similar when their individual calls were compared to the calls of Examiner Two (81.67% compared to 80.00%, respectively). However, there was a 20 percentage point difference between Blind 1 and Blind 2 when compared to the decisions generated by Examiner One (61.67% compared to 81.67%, respectively). In contrast, the comparisons between PASS decisions and the blind scorers were generally much less accurate (59.17% and 63.33%) than the comparisons between PASS and the original PDD examiners.

Table 27  
Kappa Statistics for Blind Scoring Comparisons

Comparison	Agreement %	% Chance Agreement	Value	p
Blind 1 vs. Blind 2	69.17	36.50	0.514	.0000
Examiner One				
Blind 1 vs. Examiner One	61.67	36.67	0.395	.0000
Blind 2 vs. Examiner One	81.67	41.33	0.688	.0000
Examiner Two				
Blind 1 vs. Examiner Two	81.67	41.33	0.688	.0000
Blind 2 vs. Examiner Two	80.00	38.50	0.675	.0000
PASS				
Blind 1 vs. PASS	59.17	34.15	0.380	.0000
Blind 2 vs. PASS	63.33	35.67	0.430	.0000

Note. PASS = Polygraph Automated Scoring System.

PDD Examiner/Blind Scorer Decision vs. PASS Decision. As mentioned in the introduction section, accuracy estimates for the algorithm were based on PASS having generated the same decision for a case as had been generated by the original examiner and two other examiners. Out of 120 examinations collected during this study, the original examiner and both blind scorers arrived at the same decision for 73 cases. There were 26 NDI calls, 45 DI calls and 2 INC calls. Table 28 shows that PASS arrived at a call of NDI for 100.0% of the cases labeled NDI by the group of PDD examiners. In truth, 23 of those 26 cases had been programmed innocent. The remaining three individuals had committed the theft.

For the 45 cases judged as DI by the examiners, PASS agreed with their DI call on 64.4% of the cases. Another 11.1% of the examinations were found to be NDI and the INC rate was 24.4%. Thirty-seven of the 45 subjects had been programmed guilty and the remaining 8 were innocent of committing the crime.

Of the two INC calls generated by the original examiners and blind scorers, PASS also called one exam inconclusive and the other it called NDI. In fact, one case had been programmed innocent. The other case had been programmed guilty, however this was the case which PASS labeled NDI.

Table 28

Percentage of Agreement Between PASS Decision and the Consensus Decision of the PDD Examiner and Two Blind Scorers (N = 73)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
NDI ( <u>n</u> = 26)				
Agreed	100.0	(26)	100.0	(26)
Disagreed (FP)	0.0	(0)	0.0	(0)
Inconclusive	0.0	(0)	--	--
DI ( <u>n</u> = 45)				
Agreed	64.4	(29)	85.3	(29)
Disagreed (FN)	11.1	(5)	14.7	(5)
Inconclusive	24.4	(11)	--	--
INC ( <u>n</u> = 2)				
Agreed	50.0	(1)	50.0	(1)
Disagreed	50.0	(1)	50.0	(1)
Inconclusive	0.0	(0)	--	--

Note. DI = deception indicated; FN = false negative; FP = false positive; INC = inconclusive; NDI = no deception indicated; PASS = Polygraph Automated Scoring System; PDD = psychophysiological detection of deception.

PASS/PDD Examiner/Blind Scorer Decision vs. Ground Truth.

There were 56 cases where PASS, the original examiners and both blind scorers agreed on the decision. Of the 27 cases which had been programmed innocent, PASS and the group of examiners correctly identified 23 (85.2%) of them as NDI (see Table 29). There were 3 (11.1%) cases incorrectly identified as DI and 1 case for which PASS and the examiners were unable to make a definitive call.

For the 29 cases which had been programmed guilty, PASS and the group of examiners correctly identified 26 (89.7%) of them as DI. The remaining three cases (10.0%) were incorrectly labeled NDI.

Table 29

Percentage of Agreement Between Ground Truth and the Consensus Decision of PASS, the PDD Examiner and Two Blind Scorers (N = 56)

Decision	Overall		Inconclusives Eliminated	
	%	<u>n</u>	%	<u>n</u>
Innocent ( <u>n</u> = 27)				
Agreed	85.2	(23)	88.5	(23)
Disagreed (FP)	11.1	(3)	11.5	(3)
Inconclusive	3.7	(1)	--	--
Guilty ( <u>n</u> = 29)				
Agreed	89.7	(26)	89.7	(26)
Disagreed (FN)	10.3	(3)	10.3	(3)
Inconclusive	0.0	(0)	--	--

Note. FN = false negative; FP = false positive; PASS = Polygraph Automated Scoring System; PDD = psychophysiological detection of deception.

#### Accuracy Compared Over Time

PDD Examiner Accuracy. The study design required the PDD examiners to conduct three exams per day for the same crime, throughout the entire length of the study. An analysis was performed to determine whether PDD examiner accuracy for detecting innocent and guilty individuals was affected over time. Table 30 shows that when analyzed in blocks of fifteen, both Examiner One and Examiner Two experienced an upward increase in accuracy for the first 45 exams included in the data set, (60.0% compared to 87.0%, 67.0% compared to 80.0%, respectively). Performance for both examiners dropped 20 percentage points for the final fifteen exams when compared to the preceding block of exams.

Despite a change in control question tests (CQTs) (Examiner One--DLC to PLC; and Examiner Two--PLC to DLC) at the midway point, both PDD examiners experienced an increase in accuracy for the third quarter of exams. By the time each PDD examiner began the new CQT both had collected over 40 exams, including those used in the analysis as well as those which had been excluded (i.e., pre-study procedural refinement exams and exams referred to in the subject section under Excluded Data).

Table 30  
PDD Examiner Accuracy Compared Over Time

Analyzed Exams	Examiner One		Examiner Two	
	Accuracy %	Cumulative Exams <sup>a</sup>	Accuracy %	Cumulative Exams <sup>b</sup>
1st 15	60.0	24	67.0	24
2nd 15	80.0	45	73.0	41
3rd 15	87.0	64	80.0	62
4th 15	67.0	81	60.0	80

Note. CQTs = control question test; PDD = psychophysiological detection of deception.

<sup>a</sup> cumulative total includes 4 mid-study pilot exams and 9 excluded exams in addition to 8 pre-study pilot exams. <sup>b</sup> cumulative total includes 4 mid-study pilot exams and 7 excluded exams in addition to 9 pre-study pilot exams.

PASS Accuracy. Table 31 shows a comparison of accuracy over time generated for PASS. The data are again grouped in blocks of fifteen and are also categorized as exams conducted by either Examiner One or Examiner Two. PASS's accuracy on the exams collected by Examiner One ranged from 53.3% to 66.7%, and for the exams collected by Examiner Two the range was 46.7% to 80.0%.

PASS did not show the same pattern of improvement from block to block as did the PDD examiners. Instead, the pattern generated by PASS across the blocks of exams was down-up-down-up, i.e., a decrease in accuracy on the second block of exams, a comparative increase on the third block and then another decrease on the final block of fifteen exams. PASS uses only the physiological data collected by the Axciton when generating its probability of deception. Therefore, the differences between the examiners' patterns of accuracy and the algorithm's patterns of accuracy suggests that the examiners may have been using other cues (e.g., behavioral, verbal, etc.), in addition to the physiological tracing in order to make their decisions.

Table 31  
PASS Accuracy Compared Over Time

Analyzed Exams	Examiner One Exams	Examiner Two Exams
	Accuracy %	Accuracy %
1st 15	67.0	67.0
2nd 15	53.0	60.0
3rd 15	67.0	80.0
4th 15	60.0	47.0

Note. CQTs = control question test; PASS = Polygraph Automated Scoring System.

PASS and PDD Examiner Accuracy. Table 32 and Table 33 show comparisons of PASS accuracy with the rates of accuracy for each examiner broken down by subject programming, either innocent or guilty. As pointed out in other analyses, PASS was generally more accurate at identifying innocent individuals rather than those who had committed the theft. PDD examiners were more accurate at detecting those individuals guilty of committing the crime.

Table 32  
PASS and Examiner One Accuracy Compared Over Time  
--Programmed Innocent and Programmed Guilty Cases  
(N = 60)

Analyzed Exams	Examiner One	PASS
	Accuracy %	Accuracy %
Innocent ( <u>n</u> = 30)		
1st 15 (8)	38.0	75.0
2nd 15 (7)	57.0	57.1
3rd 15 (8)	75.0	62.5
4th 15 (7)	71.0	71.4
Guilty ( <u>n</u> = 30)		
1st 15 (7)	86.0	57.1
2nd 15 (8)	100.0	50.0
3rd 15 (7)	100.0	71.4
4th 15 (8)	63.0	50.0

Note. PASS = Polygraph Automated Scoring System.



Table 33  
PASS and Examiner Two Accuracy Compared Over Time--  
Programmed Innocent and Programmed Guilty Cases  
(N = 60)

Analyzed Exams	Examiner Two		PASS
	Accuracy %		Accuracy %
Innocent ( <u>n</u> = 30)			
1st 15 (8)		63.0	87.5
2nd 15 (7)		85.0	85.7
3rd 15 (8)		63.0	87.5
4th 15 (7)		43.0	42.9
Guilty ( <u>n</u> = 30)			
1st 15 (7)		71.0	42.9
2nd 15 (8)		63.0	37.5
3rd 15 (7)		100.0	71.4
4th 15 (8)		75.0	50.0

Note. PASS = Polygraph Automated Scoring System.

## Discussion

### Accuracy

The accuracy of the Polygraph Automated Scoring System (PASS), Version 2.0 prototype software was assessed during this study using data collected under a mock crime scenario paradigm. PASS's rates of agreement and disagreement with ground truth were examined, and the same evaluations were made for the decisions of the psychophysiological detection of deception (PDD) examiners who collected the data. The ability to make direct comparisons between PASS and PDD examiner accuracy was instrumental in the assessment process, due to the manner in which PASS was developed and validated.

The PDD examiners in this study had an overall accuracy rate of 71.7% when compared to ground truth. The overall rate of accuracy generated by PASS was 62.5%. Even with the inconclusive (INC) decisions eliminated, the recomputed accuracy rate for both PASS and the PDD examiners was only adjusted to 79.0% and 79.6%, respectively. (PASS had a higher INC rate--20.8% compared to 10.0%--which accounts for the greater increase in accuracy.) Despite easily discernable differences among the rates of agreement and disagreement when compared to ground truth, a comparison of the overall accuracy level generated by PASS and the PDD examiners was not statistically significant. This lack of association had been hypothesized, however accuracy figures for both PASS and the PDD examiners were lower than expected.

Statistical significance was achieved on a number of occasions in an unpredicted area--that of innocent and guilty comparisons. With INC decisions eliminated from the computation, the proportion of innocent versus guilty individuals correctly identified by PASS was statistically significant.

In fact, there was an overriding trend in the data which indicated that PASS was more accurate in identifying innocent individuals while the PDD examiners were more accurate in identifying guilty individuals. This was illustrated in the overall innocent/guilty breakout tables. PASS had an accuracy rate of 71.7% on innocent cases compared to the PDD examiner rate of 61.7%. In contrast, the PDD examiner accuracy rate on the guilty cases was 81.7% compared to the rate of 53.3% generated by PASS. Only, the comparison between PASS and the PDD examiners' level of accuracy on the guilty cases was statistically significant, however.

As reported in the results section, the use of the directed lie control (DLC) and the probable lie control (PLC) techniques made no apparent difference, statistically, in the overall accuracy of either PASS or the PDD examiners. That had been an area of concern, due to the implementation of an untried pretesting procedure for the DLC. The comparison between test formats (DLC/PLC) for both innocent cases and for guilty cases was not statistically significant. There was, however, an indication of association between PASS decisions generated for innocent and guilty cases when using the PLC. (Note: The results of proportion tests calculated during a separate analysis did indicate that several comparisons between PDD examiner decisions made using the PLC and DLC test formats approach significance [ $.05 < p < .1$ ]. These decisions include the: inconclusive rate for all subjects; inconclusive rate for innocent subjects; and correct decision rates for the innocent and guilty subjects. While not conclusive, these results suggest that accuracy rates may differ between the PLC and the DLC test formats, when using the pretest procedures employed in this study.)

A similar within test format association on the DLC was obtained for the PDD examiners. The PDD examiners were very accurate (90.0%) at detecting the guilty subjects using the DLC. In addition, both the INC (3.3%) and the false negative (FN) rate (6.7%) were relatively low. Clearing innocent subjects using the experimental version of the DLC was no better than 50.0%. It is unclear whether the disparity in PDD examiner accuracy on the DLC innocent and guilty cases is a result of the pretest modifications. Further analysis will have to be done, apart from this evaluation geared at assessing the effectiveness of PASS.

In terms of percentage, the PDD examiners had higher overall accuracy rates and lower INC rates on both the DLC and PLC when compared to PASS. The rate of disagreement (composed of FN and

false positive [FP] decisions) was essentially comparable for both PASS and the PDD examiners. Even with the INC decisions eliminated the PDD examiners had comparable or higher accuracy rates than PASS.

#### Interrater Agreement

The comparative differences between interrater agreement were pronounced. Test results for the blind scorer decisions compared to each other, to those of the original examiner and to the decisions generated by PASS were all statistically significant. In addition, Blind 1 and Blind 2 scoring accuracy rates were 60.8% and 65.0%, respectively, when compared to ground truth. Blind scoring of examinations is commonly less accurate than scoring performed by the original examiner, and this study was no exception. Though the dissimilarity between the blind scorers and the original examiners were not excessively large, the low overall accuracy rate compared to ground truth was disappointing.

Rate of agreement between the individual blind scorers and the original examiners was much higher than the ground truth comparison. With only one exception (61.67%), the agreement figures were all approximately 80.0%. For some inexplicable reason, Blind 1 was 20 percentage points less accurate on his evaluation of Examiner One's examinations when compared to his evaluation of Examiner Two exams. There was also the same degree of disparity between his rate of accuracy and that of the other blind examiner for the same set of exams.

Though it does not explain the above mentioned discrepancy in Blind 1 scoring, there is an important note regarding blind scorer qualifications. After the fact, it was learned that the agencies which employ the blind scorers routinely use a three position scale (-1, 0, +1) as opposed to a seven position scale when scoring exams. The agencies of both original PDD examiners use the seven position scale as a matter of course. Though both blind scorers had been trained on and had used the seven position scale there are indications in the data that they scored more conservatively, than the original PDD examiners (i.e., awarding fewer  $\pm 2$ 's and  $\pm 3$ 's).

An observation which could account for low rates of interrater agreement, in this and other large scale research projects, has to do with the circumstances under which most blind scoring is conducted. Usually the tasking is in addition to the individual's normally assigned duties. In the case of much research reporting, the examiner is commonly given a short suspense for completion, as well. This may cause the blind scorer to rush through the assignment and thereby lower the rate of accuracy.

In this study there were 120 sets of exams (three tests per exam) scored by each blind examiner. Both examiners completed the tasking in under seven days. That equates to over 17 exams per eight hour day. Degradations in individual human performance have prompted some industries, such as aircrew training, to pass guidelines restricting excessive demands on performance within a duty day. No such guideline exists in the quality control (QC) area of the PDD field.

#### Defense of the Findings

With lower than expected accuracy rates for both PASS and the PDD examiners, the data was scrutinized for possible contributing factors which could account for the discouraging performance. Areas which were examined included: (a) affect of medications, (b) prior criminal record of examinee, (c) strength of the mock crime scenario, and (d) examiner performance decrements over time. A fifth area specific to PASS accuracy involved the a priori decision to use the results generated on the second run of PASS when PASS and PDD examiner decisions were in conflict. As mentioned in the scoring criteria section of this report, when the PASS decision did not match the examiner decision on the first run of the data, the second run which reflected examiner artifact editing was used.

Effect of Medications. Eleven subjects (9.2%) in this study indicated the use of prescription or over the counter drugs prior to reporting for the examination. Six of those individuals were correctly identified as programmed (four innocent and two guilty) by both PASS and the PDD examiners. The PDD examiners correctly identified three additional programmed guilty subjects (all were labeled INC by PASS). PASS correctly identified one additional innocent subject (labeled INC by the PDD examiner). This suggests that the known use of medications had little or no bearing on the overall accuracy rates.

Prior Criminal Record of Examinee. Critics of mock crime data contend that the commission of a "crime" for which there is no threat of recrimination is overshadowed by real world circumstances in which the individual may have been involved. Twenty-two subjects (18.3%) admitted to having prior arrest records and/or prison terms for offenses such as larceny, rape and murder. Of those 22 subjects, both PASS and the PDD examiners correctly identified 11 of them as programmed (5 innocent and 6 guilty). The PDD examiners correctly identified four additional individuals as guilty (PASS labeled one as NDI and three as INC). PASS identified one additional innocent subject (labeled INC by the PDD examiner).

Though the calls of PASS and the PDD examiners were predominately accurate, the effect which prior criminal activity may have had on the accuracy cited in this study is not easily

discounted. It is unclear as to whether the programmed guilty subjects were "detected" for the theft of money from the Country Store or for some underlying reason. Additionally, the analysis of this factor may be incomplete for there is reason to believe that not all individuals were forthcoming regarding their past criminal history (e.g., during the pretest interview one subject was asked why he had left his former job as a cook. He replied that he had been paroled from prison. The individual made no mention of a criminal record in his demographic form.).

Strength of the Mock Crime Scenario. In the scenario section of this report it was explained that a deliberate intruder was used to heighten the arousal level of the subjects tasked with committing the theft. Anecdotal analysis of the data indicates that the use of the deliberate intruder may have been more effective for some subjects than for others. Numerical evaluations for the programmed guilty individuals, who were judged DI by the PDD examiners, ranged from minus (-) 3 to minus (-) 45. Though a number responded to either, or both the strong and weak relevant questions, the majority of guilty individuals were labeled DI on the basis of a minus (-) 3 score on the evidence connecting question (i.e., "Do you know where any of that stolen money is now?"). Since all guilty subjects had the money with them in the PDD exam suite this finding is not surprising.

Within the realm of this study it is impossible to account for the disparity in numerical evaluations. To the extent possible, considering human variability, all guilty subjects were programmed in exactly the same manner (i.e., tone, inflection, level of enthusiasm shown by the scenario setter, etc.)

Examiner Performance Decrements Over Time. As shown in the results section, there was a marked difference in PDD examiner accuracy when exam decisions were analyzed in sequential blocks of 15. One can only speculate as to why the accuracy rates were arrayed in such a way. Both PDD examiners appeared to be comfortable with the process prior to beginning formal data collection, since each had conducted at least eight prestudy exams.

The increase in accuracy between the first and second block of exams could potentially be accounted for by having settled into the daily routine and by having eliminated any residual trepidation regarding the procedures. The increase between the second and third block of exams may have been due to renewed interest and vigilance, as a result of the control question test (CQT) switch from DLC to PLC and vice versa. By that point, both PDD examiners had collected over 40 exams for the same crime using the same technique.

The dramatic drop between the third and fourth block of exams can probably be attributed to a much anticipated end to the study and also to the sheer boredom with having run the same exam, three times a day for almost five weeks. It would be very unlikely that a field PDD examiner would be required to run such a high number of exams for the same crime. In that regard, the downward shift in the level of accuracy experienced in this study would not be an issue in a field investigation. Also, had the level of accuracy for the fourth block of exams at least remained the same as for the third block of exams, the overall accuracy rate of the PDD examiners would certainly have been raised several percentage points.

In keeping with that thought, it is appropriate to say that PASS is only as accurate as the collected examination enables it to be. If the control questions have not been properly set by the PDD examiner due to distraction or boredom, PASS cannot be held accountable for its inability to detect differences in physiological responding. However, PASS accuracy on the four blocks of exams did not follow the same pattern as that of the PDD examiners. Though PASS and the PDD examiners had similar rates of accuracy on the first block of exams, PASS accuracy dropped on the second block of exams, some 27 and 13 percentage points below the increased PDD examiner accuracy rates for the same set of tests. Similar results occurred on the third and fourth blocks of exams.

Artifact Editing. As mentioned earlier, PASS both enables and occasionally requires subjective manipulation of the data in the area of artifact editing. There were 17 cases which resulted in a conflict between the PDD examiner decision and the PASS decision. Had the first run data been used in the study it would have resulted in four additional correct decisions by PASS, (three NDI and one DI). For the remainder of the 17 cases PASS would have been credited with 2 FP, 4 FN, and 7 INC (5 programmed innocent and 2 programmed guilty).

This review of the factors suspected in adversely affecting accuracy rates of both PASS and the PDD examiners revealed that adjustments in only one area could have conclusively and substantially increased overall levels of accuracy. That area was examiner performance decrements over time.

#### PASS Accuracy via APL Assessment Criteria

In keeping with the "examiner consensus" criterion used by APL to develop and validate PASS (outlined in the introduction section), a direct comparison was made between PASS decisions and the collectively agreed upon decisions of the original PDD examiner and two blind scorers. There were 73 cases for which the PDD examiners and blind scorers arrived at the same call. PASS agreed with all 26 NDI calls (100.%) made by the group of examiners. That finding corresponds to the information in Table

3, where APL cites PASS accuracy at 93.5% with INC decisions and at 100.0% without INC decisions for 91 confirmed cases with known ground truth. (In actuality in this comparison, PASS and the PDD examiners generated 23 correct NDI calls and 3 FN decisions.)

However, the DI findings in this study were not comparable to those of APL during algorithm development. Of the 45 cases judged as DI by the PDD examiners and the blind scorers, PASS agreed with the calls on only 64.4% of the cases. (PDD examiners generated 37 correct DI calls and 8 FP; PASS generated 29 correct DI calls, 5 FP and 11 INC.) It was later shown that the group of examiners had been 90.2% accurate on their DI decisions when compared to ground truth, further eroding PASS's level of accuracy on the DI decision cases.

Finally, a comparison was made between ground truth and the 53 cases on which PASS, the PDD examiners and the blind scorers agreed on the call. Accuracy assessments in this instance were 85.2% for programmed innocent individuals and 89.7% for programmed guilty individuals. This indicates that a mutual decision reached by PASS, the original examiner and two blind scorers is far more likely to be accurate than a decision reached by PASS alone, or in conjunction with the original examiner. However, even with the agreement of PASS and all the examiners in this study the overall accuracy rates for innocent and guilty individuals are below those achieved by APL during algorithm development.

### Conclusions

The Polygraph Automated Scoring System (PASS), Version 2.0, is the prototype for a user friendly software package. Using mock crime data obtained during this research effort, it provided moderate levels of accuracy when coupled with the decisions of the original examiner and two blind scorers. It was less accurate when used alone or in conjunction with the original examiner decision. PASS was biased toward clearing the innocent individual rather than detecting the guilty individual. In this study the reverse statement was true of the participating PDD examiners. Both observations were found to be statistically significant.

According to the Applied Physics Laboratory (APL), one of the primary advantages PASS offers the psychophysiological detection of deception (PDD) field is scoring consistency. Due to the fact that PASS is a computer-based system, it does offer scoring consistency--but only when the same artifacts are edited with consistency by all the scoring examiners. In the Version 2.0 configuration, modest variability in subjective editing can and does impact the resulting PASS outcome. Differences in interrater agreement observed in this study point to a need for a truly consistent method for evaluating PDD examinations.

Future directions for research stemming from this study include an indepth look at the differences between accuracy rates achieved when using the directed lie control (DLC)--experimental version. Understanding why the test was so much more accurate in detecting the guilty individuals could be used in future test development efforts. Also, pinpointing the reasons behind the wide margin of numerical scores on the guilty cases would benefit PDD research in the area of more effective mock scenario development.

A final note regarding PASS effectiveness is the planned evaluation of current and future releases of the PASS software. Versions released since the conduct of this study reportedly offer revamped data handling, increased accuracy and improved user friendliness. Using the data set collected during this study, each new software release will be evaluated in order to assess whether the system improvements correct previously identified scoring errors.



## References

- Capps, M. (1993, January). Polygraph automated scoring system (PASS), version 2.0. Paper presented at the Department of Defense Polygraph Institute training program on PASS Operations, Fort McClellan, AL.
- Crunch Software Corporation. (1991). Crunch, version 4.0 [Computer program]. Oakland, CA.
- Department of Defense Polygraph Institute. (1992). Chart Evaluation. (Available from Department of Defense Polygraph Institute, Building 3195, Fort McClellan, AL).
- Department of Defense Polygraph Institute. (1992, August). Zone Comparison Technique (ZCT). (Available from Department of Defense Polygraph Institute, Building 3195, Fort McClellan, AL).
- The Johns Hopkins University Applied Physics Laboratory. (1993a). Polygraph automated scoring system (PASS)--user's guide, version 2.0 [Computer program manual]. (SSD/POR-93-6799). Laurel, MD.
- The Johns Hopkins University Applied Physics Laboratory. (1993b). Polygraph automated scoring system (PASS), version 2.0 [Computer program]. Laurel, MD.
- Nunnally, J. C. (1978). Psychometric Theory (2nd ed.). New York: McGraw-Hill.
- Office of Technology Assessment. (1983). Scientific validity of polygraph testing: A research review and evaluation--a technical memorandum (Report No. OTA-TM-H-15). Washington, DC: U.S. Government Printing Office.

## Appendix A

### Chart Evaluation Criteria

#### **PNEUMOGRAPH**

- o Change in rhythm and regularity
- o Change in amplitude or volume
- o Change in inhalation-exhalation strokes
- o Change in baseline
- o Loss of baseline
- o Notched or serrated strokes
- o Hyperventilation
- o Suppression
- o Holding or blocking (apnea)

#### **GALVANOGRAPH**

- o Vertical rise at point of deception
- o Complex response
- o Long duration and/or degree of response at stimulus
- o Plunging tracing at stimulus

#### **CARDIO**

- o Increase and decrease in blood volume
- o Increase only in blood volume
- o Decrease only in blood volume
- o Increase in pulse rate
- o Decrease in pulse rate
- o Increase in amplitude (change in blood pressure)
- o Decrease in amplitude (change in blood pressure)
- o Change in position or disappearance of dicrotic notch
- o Extra systoles (pre-ventricular contraction)

## Appendix B

### "Programmed Innocent"

#### SCRIPT

Today there was a mock crime committed. The crime was a "theft" of money from one of the rooms in the main DoDPI building. Since you did not have any part in that crime, you obviously do not know any of the details of that crime. However, because you were seen near the area where the "theft" occurred you are now a suspect in this case.

In a little while you will be introduced to a PDD examiner, who will ask you to take a PDD examination. I would like you to go with him and take the test. He will ask you what you know about the "theft". Simply tell him the truth. Tell him that you were told that a "theft" of money had occurred today, but that you have no involvement in the crime and that you have no knowledge of any of the details. In every other way I would like you to be as cooperative as possible and do your best to follow all the PDD examiner's instructions.

If you feel uncomfortable about anything, tell the PDD examiner that you would like to talk to your "lawyer." That will be the code that something is wrong and that you need to talk to me. Thank you again for your cooperation.

## Appendix C

### "Programmed Guilty"

#### SCRIPT

**SCENARIO SETTER:** Today YOU are going to commit a "theft" of money from the Country Store here at DoDPI. It is very important that you follow all my instructions to the letter. There are a number of details which you must remember, for you will be tested later. Are there any questions before we go to the room?

(Subject and scenario setter go to the "crime" scene.)

(Subject's First Name), this is the DoDPI Country Store. See all that candy, cookies, popcorn and other snacks? This stuff sits out here on the table all the time so people can come in and buy what they want. The Country Store operates on the honor system. When someone takes something, they are supposed to put the money for it in this box.

(Scenario setter place hands on box.)

You can see how easy it would be for someone who is dishonest to take something and not pay for it. You can also see that if someone was really dishonest they could just reach into that box and steal all the money. Today you are going to be dishonest and steal all of that money.

(Scenario setter opens the box so the subject can see the money.)

There's a lot of money in there, too. I'll bet there's at least \$20 in there. A lot of honest people have bought something to eat and put money in there, but you're going to steal it all and keep it for yourself. O.K.? You stand right here in front of the table.

(Scenario setter positions the subject in front of the table.)

**SCENARIO SETTER:** I'm going to leave you alone in here. When I disappear through that doorway (point to lounge door leading to hallway), I want you to open that money box and take out all of the paper money. Stand here and quickly count it. Then hide it somewhere on your body, in a pocket or in your shoe or wherever. When you've hidden it meet me in the hallway outside that door. Understand? Wait until I leave, take the money out of the box, count it, hide it and then meet me outside the door.

For the sake of the project, it's very important that you don't get caught stealing the money. Understand? There are people coming in and out of here all the time. If someone comes in before you get the money counted and hidden, you'll have to act like nothing is going on. If someone catches you, act like you're buying something and are just making change. Close the box, walk over to this table, pick up a candy bar and then meet me in the hallway outside that door. Any questions?

(Scenario setter leaves the lounge through door leading to hallway and stands just outside the door.)

**DELIBERATE  
INTRUDER:**

(Deliberate intruder stands just inside the library doorway, and watches for scenario setter to leave. After counting to 10 enter the lounge to surprise the subject. Speak to the subject as you pass and go to the table as if to pick out a snack to buy. Stay in the room until the subject leaves to join the scenario setter in the hallway.)

**SCENARIO SETTER:** (Lead the subject down the hall to the forms closet. Partially block the doorway. Make sure the subject has the money. If he/she didn't finish counting it, have them count it and then hide it on their body. Get the candy bar back so it can be returned to the Country Store. Lead the subject down the hallway and into the holding area. Sit at one of the tables and have the subject complete the "Crime Scene" Scenario Questionnaire. When he/she has completed the questionnaire review what the subject has just done.)

**SCENARIO SETTER:**

Today you committed a "theft" from the DoDPI Country Store. You stole \$124 and almost got caught doing it. In a little while a PDD examiner will be asking you to take a PDD exam. I would like you to go with him and take the test. He will ask you what you know about the "theft." Do not, under any circumstances tell him what you did today. Even though you're guilty of stealing the money and you will in fact still have it hidden on you when you go in to take the PDD examination in a few minutes, you have to make the PDD examiner believe that you don't know anything about the "theft."

Simply tell him that you were told that a "theft" had been committed but that you have no involvement in the crime and that you have no knowledge of any of the details. In every other way I would like you to be as cooperative as possible and do your best to follow all the PDD examiner's instructions, but DO NOT confess to having any knowledge or involvement in the crime.

If you have any questions or feel uncomfortable about anything, tell the PDD examiner that you would like to talk your "lawyer." That will be the code that something is wrong. Thank you again for your assistance.

Appendix D

"Crime" Scene Scenario

**QUESTIONNAIRE**

You have just participated in a mock crime scenario. Please answer for the following questions. Take your time. Do not discuss the contents of this statement with anyone other than DoDPI personnel.

1. Where did the Scenario Setter take you to "commit" your crime?

- ☐ Country Store
- ☐ Restroom
- ☐ Trailer
- ☐ Copier Room

2. What kind of "crime" did the Scenario Setter ask you to commit?

- ☐ Theft of government property
- ☐ Theft of money
- ☐ Theft of classified documents
- ☐ Theft of personal property

3. What item did you steal?

- ☐ Woman's purse
- ☐ Top Secret document
- ☐ \$124.00 in cash
- ☐ Army Jeep

4. How did you feel while committing the "theft"?

- ☐ Nervous
- ☐ Cheerful
- ☐ Not affected
- ☐ Other (please specify)\_\_\_\_\_

5. Did anyone see you steal the item?

- ☐ Yes
- ☐ No

6. Other than the Scenario Setter, did anyone see the item in your possession?

- ☐ Yes
- ☐ No

7. Where did you hide the item you stole?

---

8. Do you feel that you are guilty of committing a "theft" today?

☐ Yes

☐ No

9. Do you think you can hide your "guilt" from the PDD Examiner?

☐ Yes

☐ No

10. Do you think that you will pass your PDD examination today?

☐ Yes

☐ No

Please return the completed form to the Scenario Setter.

Thank you.



## Appendix E

### Welcome and General Briefing

Hi, my name is Joan Blackwell. Welcome to the Department of Defense Polygraph Institute, also known as DoDPI. This may be the first time you have been at DoDPI so we would like to provide you with some information concerning the purpose for being here today. We hope that you will find your time here to be enjoyable and educational.

Allow me to start by defining a term that you will see and hear several times today: PDD examination. PDD stands for psychophysiological detection of deception; you may have heard it referred to as "taking a polygraph test." That's what we are going to ask you to do today.

We have two missions here at DoDPI. To begin with, we are one of only two schools in the Federal Government that trains PDD or polygraph examiners. We train all the Department of Defense PDD examiners and most of the other federal agencies, such as the Secret Service, the Drug Enforcement Agency (DEA), the Federal Bureau of Investigation (FBI), etc. The other part of our mission here is to conduct research. We test all the new and existing PDD procedures for accuracy and utility. It is in that capacity that we are asking for your assistance today.

One of the ways that we test a particular procedure for accuracy is to ask people like you to commit a make believe crime. We then give you a PDD examination to see if we can determine that you did commit that crime.

Of course if everyone we tested was guilty, then we would not have a very good experiment. That's why we also test some people who did not commit the mock crime and are therefore innocent. Today we may make you part of an innocent group or part of a guilty group. In either case it is very important that you do exactly as instructed before, during, and after your PDD examination.

Just as a brief explanation, during a PDD examination a trained investigator will talk to you for a few minutes about your background, he will tell you what he knows about a crime committed here at DoDPI, and will ask you what you know about that crime. He will then attach sensors to your body to record your breathing pattern, heart rate and sweat gland activity. Use of these sensors will determine whether you are being truthful when asked questions about a particular incident. If you choose to participate in the project today the PDD examiner will provide you with a more in-depth and detailed explanation about the instrumentation and procedures used.

Your participation is completely voluntary. I would like to assure you in advance that we will not ask you any embarrassing questions or make you do anything that you are uncomfortable doing. However, if you have reservations about participating in a mock crime and then lying about it, or if you doubt that you will be able to cooperate fully for any reason, please notify me or any other DoDPI staff member. If you have any questions at any time during the day, please feel free to ask any of the DoDPI staff. Are there any questions at this time?

**(ANSWER QUESTIONS)**

Has anyone here ever taken a PDD or polygraph examination before?

**(ELIMINATE SUBJECTS WITH PRIOR PDD EXPERIENCE)**

What I would like to do now is provide you with a packet of information about this project. The packet includes a Project Briefing document, a Volunteer Agreement Affidavit, and a Background Information Form. Please read over the Project Briefing document now. If, after doing so, you choose to participate in the project, go ahead and read and complete the other two forms. I'll come around and witness your signature when you're finished.

**(ALLOW TIME FOR FORM COMPLETION)**

Now that we've completed the required paperwork, I'll turn you over to \_\_\_\_\_, who will escort you throughout the day.

## Appendix F

### Project Briefing

1. **PROJECT TITLE:** "An Evaluation of the Effectiveness of the Polygraph Automated Scoring System (PASS) in Detecting Deception in a Mock Crime Analog Study."
2. **PRINCIPAL INVESTIGATOR:** N. Joan Blackwell  
DoDPI Research Psychologist
3. **DISCUSSION:** Congress has directed the DoDPI to conduct research to determine the effectiveness of psychophysiological detection of deception (PDD) examinations. (PDD examinations are commonly known as polygraph tests). Part of this mandate requires that new and existing PDD scoring procedures be tested for accuracy and reliability. You are being asked to volunteer for an investigation that will help us investigate the accuracy of a computerized PDD scoring algorithm, or evaluation technique. You may or may not be involved in a mock "theft" scenario, today. If you are designated to commit a "theft," then you will be asked to follow certain instructions from a DoDPI staff member, known as the Scenario Setter. After following those instructions, you will be asked to take a PDD examination. If you are not asked to commit a "theft," then you will be taking a brief PDD examination regarding a matter in which you will obviously have no direct involvement. You will each be instructed and tested individually.
4. **DISCOMFORTS:** Part of the PDD process requires the wearing of an inflated blood pressure cuff, which can be moderately uncomfortable. As a result, a few people find it difficult to sit still during the PDD examination. The cuff is inflated for approximately five minutes per test and each PDD examination is composed of up to five tests. The PDD examiner is sensitive to this discomfort and strives to minimize the length of time the cuff is inflated, but he needs your complete cooperation. The total length of time required for your participation in this project will be approximately 1-1 1/2 hours, however, you may be here at the DoDPI for the entire day.
5. **VIDEOTAPING:** All examinations conducted during this project will be videotaped using wall and ceiling mounted video cameras and commercial videotape recorders. The tapes collected will be maintained until completion of the operational and data analysis portions of this project are complete. At that time the video tapes will be erased and made available for re-use by research and instruction divisions.

**6. RISKS:** There are no known risks involved in this project. The research protocol has been reviewed and approved by members of a Human Use Committee (HUC) composed of various medical and military personnel external to the DoDPI.

**7. CONFIDENTIALITY OF RECORDS:** Admissions about a serious crime, or a breach of national security may be reported to the proper authorities for investigation. However, all other information you tell the PDD examiner is privileged information and will not be revealed to anyone not directly involved in the project. As a result, (in the absence of serious crime and national security related admissions), all of the videotapes, polygraph tests, score sheets, interview forms, PDD examiner work sheets, and related documents associated with your examination will be used for research purposes only. Members of the Army Surgeon General's Human Subjects Research Review Board may inspect the records of the research in their capacity as reviewing officials, but your identity will be kept confidential.

**8. YOUR RIGHTS:** You have the right to ask any questions about any aspect of your participation in the project. If problems arise in conjunction with your involvement in the project, or if you have been injured in any way as a result of the project, the person to contact is Dr. William J. Yankee, Director of the DoDPI. In the event that you do have questions or any of the above has occurred, Dr. Yankee can be reached at (205) 848-3804. Should any question arise concerning project-related injury, you may contact COL. Roland J. Weisser, Jr., Director of the Noble Army Community Hospital, Fort McClellan, Alabama, 36205, at (205) 848-2200.

**9. VOLUNTARY PARTICIPATION:** Your participation in this project is completely voluntary. If you would prefer not to participate, do **NOT** volunteer for it! Even if you initially agree to participate in the project, you may discontinue at any time without penalty or recrimination. Simply inform anyone on the DoDPI staff. If you decide to withdraw during the PDD examination itself, inform the PDD examiner and you will be returned to the waiting area where you will be debriefed and released.

## Appendix G

### Volunteer Agreement Affidavit

This form is affected by the Privacy Act of 1974.

1. AUTHORITY: 10 USC 3012, 44 USC 3101 and 10 USC 1071-1087.
2. PRINCIPAL PURPOSE: To document voluntary participation in the Defense Polygraph Institute Research Program. Your name will be used for identification.
3. ROUTINE USES: The name will be used for identification and locating purposes. Information may be furnished to Federal, State, and local agencies.
4. MANDATORY OR VOLUNTARY DISCLOSURE: Your signature is necessary if you want to be included in this research. If you do not sign, you will not be able to serve in this study.

---

### PERSONAL STATEMENT

I, \_\_\_\_\_, being at least 19 years old, do hereby volunteer to participate in a research project entitled "An Evaluation of the Effectiveness of the Polygraph Automated Scoring System (PASS) in Detecting Deception in a Mock Crime Analog Study," being conducted by the Department of Defense Polygraph Institute (DoDPI) at Fort McClellan, under the direction of N. Joan Blackwell.

1.\_\_\_\_\_I understand that I am participating in a research project to determine the utility of the PASS algorithm in scoring a criminal PDD examination.

2.\_\_\_\_\_I am aware that I will be spending approximately four (4) hours at DoDPI and that during this time I may be asked to participate in the commission of a mock "theft."

3.\_\_\_\_\_I understand that there are no known dangers or risks arising as the result of my participation in this project.

4.\_\_\_\_\_I understand that as a part of this project, I will be taking a PDD examination, during which I will be asked to sit still for several minutes at a time during up to five PDD tests, while psychophysiological measurements are being recorded from my body.

5.\_\_\_\_\_I understand that part of the PDD examination process requires the wearing of an inflated blood pressure cuff, which some people find moderately uncomfortable.

6.\_\_\_\_\_I understand that I will be videotaped during the PDD examination and that the videotape will be maintained for additional study until the project has been completed, at which time the tape will be erased and used again for research or instruction purposes.

7.\_\_\_\_I understand that I will receive no reward or benefit of any kind as the result of my participation in this project.

8.\_\_\_\_My participation, the nature, duration and purpose of the investigation and the methods by which it is to be conducted, have been thoroughly explained to me. I have been given the opportunity to ask questions concerning this project, and any such question has been answered to my satisfaction.

9.\_\_\_\_I understand that I may terminate my involvement in this project at any time and for any reason.

10.\_\_\_\_Should I have any concerns or complaints concerning this project, I understand that I may contact N. Joan Blackwell, or Dr. William J. Yankee at (205) 848-3804.

11.\_\_\_\_Should any question arise concerning my rights relating to project-related injury, I should contact COL. Roland J. Weisser, Jr., Director of the Noble Army Community Hospital, Fort McClellan, Alabama, 36205, at (205) 848-2200.

---

**SIGNATURE**

---

**WITNESS**

---

**PRINTED NAME**

---

**PRINTED NAME**

---

**DATE**

---

**DATE**



**Sleep:** How much sleep did you get last night?\_\_\_\_\_

<b>Arrest Record:</b>	<u>Month</u>	<u>Year</u>	<u>Location</u>	<u>Offense</u>

<b>Employment:</b>	<u>Month</u>	<u>Year</u>	<u>Employer</u>	<u>Position</u>

<b>Family Background:</b>	<u>Name</u>	<u>Age</u>	<u>City/State</u>	<u>Occupation</u>
Mother				
Father				
Sister(s)				
Brother(s)				
Spouse				
Children				

**Hobbies:** \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

DO NOT WRITE BELOW THIS LINE

-----

PDD EXAMINER COMMENTS



Appendix I

Directed Lie Control (DLC)  
Pretest Interview

**SCRIPT**

AGENT: Hi, my name is Special Agent \_\_\_\_\_ (Your name) \_\_\_\_\_, of the \_\_\_\_\_ (Your Agency) \_\_\_\_\_. I have been assigned to administer \_\_\_\_\_ a PDD examination to you today. I see here by the paperwork that your name is \_\_\_\_\_ (Subject's full name) \_\_\_\_\_. Tell me \_\_\_\_\_ (Subject's first name) \_\_\_\_\_, do you know why you are being administered a PDD exam today?

: *(The examinees, regardless of their programming have been told that they will be examined as possible suspects in a "theft" of money).*

AGENT: Well \_\_\_\_\_ (Subject's first name) \_\_\_\_\_, since you are a suspect in this investigation, I need to make sure that you understand that you have certain rights in this investigation.

*(At this time the suspect is informed of their rights and asked to sign a Rights Advisement Form and a Consent to Interview With Polygraph form.)*

AGENT: The next thing we are going to do is to look over the background form that you have already filled out. It is important to make sure that all the information is absolutely correct. Among other things, we need to make sure that you are physically suitable to take a PDD examination.

*(At this time the examiner goes over the subject's Background Information Form. Take no more than 5-10 minutes to do this.)*

AGENT: Now \_\_\_\_\_ (Subject's first name) \_\_\_\_\_, let me tell you a little about a polygraph instrument and how it works.

*(Briefly describe the Axciton components and discuss Fight, Flight, Freeze. Then attach the components and run a Stimulation Test. Explain that his/her body will*

*react the same way on the test about the theft if he/she is telling a lie, but to make sure that his/her body is still capable of responding you are going to ask him/her to deliberately lie to you on some questions.)*

AGENT: Now       (Subject's first name)      , you have told me that you have no involvement in the crime that was committed today. Very soon I will be using the polygraph instrument to ask you certain questions regarding this crime. To each of these questions I want you to answer truthfully. If you have no involvement in this crime and are completely truthful with me about it, then I should not see your body respond when I ask you any of these questions.

*(Review the sacrifice relevant and relevant questions.)*

On this next group of questions I'm going to ask you to tell me a deliberate lie, just like I did on the numbers test. Like I said before, I want to make sure your body is still capable of responding while I'm asking you questions. I want you to think of an incident when you stole something. I don't want you to tell me about it, I just want you to picture it in your mind. We aren't hardened criminals, and we don't mean to, but technically we all steal from time to time. Think back to when you were growing up, your mother may have sent you to the store for something and you kept the change when you got back. As adults, we steal in different ways. We might take off from work early and still get paid for a full day or maybe take home a nice pen or pencil from the office. We've all taken things that didn't belong to us, wouldn't you agree? You've done that haven't you? Well now I'm going to ask you if you've ever stolen anything and I want you to lie to me. When I ask you these next three questions I want you to think of a time when you did steal something and then lie to me about it.

*(Review the control questions. Explain that the next questions will show you how his/her body responds when telling the truth and then review the irrelevant questions. Mention the need for trust and then review the symptomatic questions.)*

I will be collecting several PDD tests and on each of the tests I will be asking the questions we just reviewed, but they will be in a different order each time. It is very important that you sit absolutely still during the test and that you not talk during the test, except to answer each of my questions. Do you have any questions? If not, then let's proceed.

*(The examiner attaches the components and proceeds to collect 3-4 ZCT tests.)*

Appendix J

Probable Lie Control (PLC)  
Pretest Interview

**SCRIPT**

AGENT: Hi, my name is Special Agent       (Your name)      , of the       (Your Agency)      . I have been assigned to administer a PDD examination to you today. I see here by the paperwork that your name is       (Subject's full name)      . Tell me       (Subject's first name)      , do you know why you are being administered a PDD exam today?

*(The examinees, regardless of their programming have been told that they will be examined as possible suspects in a "theft" of money).*

AGENT: Well       (Subject's first name)      , since you are a suspect in this investigation, I need to make sure that you understand that you have certain rights in this investigation.

*(At this time the suspect is informed of their rights and asked to sign a Rights Advisement Form and a Consent to Interview With Polygraph form.)*

AGENT: The next thing we are going to do is to look over the background form that you have already filled out. It is important to make sure that all the information is absolutely correct. Among other things, we need to make sure that you are physically suitable to take a PDD examination.

*(At this time the examiner goes over the subject's Background Information Form and sets the controls according to the ZCT Question List. Take no more than 15-20 minutes to do this).*

AGENT: Now       (Subject's first name)      , let me tell you a little about a polygraph instrument and how it works.

*(Briefly describe the Axciton components and discuss Fight, Flight, Freeze.)*

AGENT: Now (Subject's first name), you have told me that you have no involvement in the crime that was committed today. Very soon I will be using the polygraph instrument to ask you certain questions regarding this crime. To each of my questions I want you to answer truthfully. If you have no involvement in this crime and are completely truthful with me about it, then I should not see your body respond when I ask you any of the questions. If however, I do see responses to one or more of the questions, then I will of course know that you do have knowledge of or are involved in this crime. These are the questions that I will be asking you today.

*(Briefly review the questions and have the Subject answer the questions as he/she plans to answer them on the test.)*

I will be collecting several PDD tests and on each of the tests I will be asking the questions we just reviewed, but they will be in a different order each time. It is very important that you sit absolutely still during the test and that you not talk during the test, except to answer each of my questions. Do you have any questions? If not, then let's proceed.

*(The examiner attaches the components and proceeds with the Stimulation test. After briefly discussing the results of the Stimulation test with the Subject, the PDD examiner will collect 3-4 ZCT tests. There will be no talking between tests.)*

Appendix K

Rights Advisement Form

Place \_\_\_\_\_

Date \_\_\_\_\_

Time \_\_\_\_\_

Before we ask you any questions, you must understand your rights.

You have the right to remain silent.

Anything you say can be used against you in court.

You have the right to talk to a lawyer for advice before we ask you any questions and to have a lawyer with you during questioning.

If you cannot afford a lawyer, one will be appointed for you before any questioning if you wish.

If you decide to answer questions now without a lawyer present, you will still have the right to stop answering at any time. You also have the right to stop answering at any time until you talk to a lawyer.

**WAIVER OF RIGHTS**

I have read this statement of my rights and I understand what my rights are. I am willing to make a statement and answer questions. I do not want a lawyer at this time. I understand and know what I am doing. No promises or threats have been made to me and no pressure or coercion of any kind has been used against me.

Signed \_\_\_\_\_

Examiner \_\_\_\_\_

Witness \_\_\_\_\_

Appendix L

Consent to Interview With Polygraph

Place \_\_\_\_\_

Date \_\_\_\_\_

Time \_\_\_\_\_

Before we begin an examination by means of the polygraph in connection with \_\_\_\_\_

\_\_\_\_\_

you must understand your rights.

**YOUR RIGHTS**

You have the right to refuse to take the polygraph test.

If you agree to take the polygraph test, you have the right to stop the test at any time.

If you agree to take the polygraph test, you have the right to refuse to answer any individual questions.

**WAIVER AND CONSENT**

I have read this statement of my rights and I understand what my rights are. I voluntarily agree to be examined by means of the polygraph during this interview. I understand and know what I am doing. No threats or promises have been used against me to obtain my consent to the use of the polygraph. I understand that the examination room (DOES) (DOES NOT) contain an observation device and that the examination (WILL) (WILL NOT) be monitored or recorded.

Signed \_\_\_\_\_  
Examinee

\_\_\_\_\_  
Examiner

Witness \_\_\_\_\_

Appendix M  
Stimulation (numbers) Test

**Example**

- X This test is about to begin.
3. Regarding the number you wrote, did you write the number 3?
4. Did you write the number 4?
5. Did you write the number 5?
6. Did you write the number 6?
7. Did you write the number 7?
8. Did you write the number 8?
- XX This test is about to end. Please remain still while I take the instrument out of operation.



## Appendix N

### Zone Comparison Test (ZCT)

#### QUESTION LIST

- X This test is about to begin.
- I1. Are the lights on in this room?
- SR. Regarding that stolen money, do you intend to answer truthfully each question about that?
- S3. Are you completely convinced that I will not ask you a question on this test that has not already been reviewed?
- C1. Prior to 1993, did you ever steal anything?
- R1. Did you steal any of that money?
- C2. Prior to your last birthday, did you ever steal anything from  
someone who trusted you?
- R2. Did you steal any of that money from the Country Store?
- S8. Is there something else you are afraid I will ask you a question about even though I have told you I would not?
- C3. Prior to this year, did you ever steal anything from an employer?
- R3. Do you know where any of that stolen money is now?
- XX This test is about to end. Please remain still while I take the instrument out of operation.

Note: I = irrelevant question  
SR = sacrifice relevant question  
S = symptomatic question  
C = control question  
R = relevant question

## Appendix O

### Subject Debriefing Form

Now that you have completed your role in our research study, It is the desire of the entire project staff to take this opportunity to sincerely thank you for your help. Your work here today was more important than you may realize.

If the results of this study show that this scoring procedure is useful, then we may be able to provide federal agencies and police departments with a new and highly accurate way to determine whether a person has knowledge or involvement in a criminal offense.

For those of you who actually committed a mock crime today, you are assured by the staff of this institute, that you in no way violated any rule or law. The mock crime was just that, pretend.

For those of you who committed no mock crime, your role was just as important, as no polygraph procedure is useful if it cannot identify the innocent as well as the guilty.

Regardless of your role, it is our hope that nobody involved in this study has made you uncomfortable in any way. If you do have questions or concerns please bring them to the attention of your briefer or to Dr. William Yankee, Director, DoDPI.

Lastly, and most importantly, DO NOT discuss the details of this study with anyone else.

This is particularly important for those of you who have knowledge regarding our mock crime scenario. If you go back to your unit or community and tell other individuals what happened in that crime scene, then they too will have guilty knowledge. If one or more of those individuals are subsequently asked to participate in this study as "innocent" people, the guilty knowledge that YOU gave them will cause false results and seriously damage this project.